Provably Stable Learning Control of Linear Dynamics With Multiplicative Noise

Peter Coppens and Panagiotis Patrinos

Abstract—Control of linear dynamics with multiplicative noise naturally introduces robustness against dynamical uncertainty. Moreover, many physical systems are subject to multiplicative disturbances. In this work, we show how these dynamics can be identified from state trajectories. The least-squares scheme enables the exploitation of prior information and comes with practical data-driven confidence bounds and sample complexity guarantees. We complement this scheme with an associated control synthesis procedure for linear quadratic regulator (LQR) that robustifies against distributional uncertainty, guarantees stability with high probability, and converges to the true optimum at a rate inversely proportional to the sample count. Throughout, we exploit the underlying multilinear problem structure through tensor algebra and completely positive operators. The scheme is validated through numerical experiments.

Index Terms—Identification for control, statistical learning, stochastic optimal control, uncertain systems.

I. INTRODUCTION

RECENTLY, the control community has gained renewed interest in learning control [1]. This can be viewed as a reaction to the practical success of machine learning—specifically reinforcement learning (RL)—when applied to control dynamical systems. An opportunity arises in the lack of theoretical guarantees, particularly those related to reliability and safety. In that regard, control theory has much to offer [1], [2]. This work focuses on stability, specifically [3, §2.1.3].

One approach toward learning with guarantees is to focus on specific classes of dynamics. The linear quadratic regulator (LQR) [4, §4.1] has proven a valuable subject for such investigations. We categorize two approaches: the model-free approach takes an RL algorithm and applies it to LQR, proving convergence, stability, etc. [5], [6], [7]; and the model-based approach develops learning control schemes specialized to LQR with guarantees [8], [9], [10].

The latter, model-based approach [11, p. 360], [12, p. 537], [4, §6.1], involves first identifying the dynamics before synthesizing a controller. As suggested in [1] under the term coarse-id learning, integrating model-based approaches with statistical learning theory yields verifiably safe controllers [8]. We also present such a coarse-id method here.

Specifically, we present a distributionally robust (DR) generalization [13] of this coarse-id framework for deterministic LQR, tailored to linear dynamics with multiplicative noise. It incorporates a least-squares estimator of the dynamics paired with tight error bounds and accompanied by a control synthesis scheme. In this manner, we expand upon our prior work [14] and interpolate between robust and stochastic control, becoming less conservative as data is gathered while guaranteeing stability with high probability.

LQR under multiplicative noise was first considered by Wonham [15], who developed a generalized Riccati equation. Similarly, many recent developments in learning LQR are generalizable to multiplicative noise [14], [16], [17], [18]. This is interesting for two reasons.

First, multiplicative noise affects numerous real-world systems. These models occur frequently when disturbances enter a model through the parameters (e.g., through vibrations, nonlinear effects in spring systems, thermal, and aerodynamic influences, etc.). A list of examples with references is provided in [19, §1.9] with concrete applications in aerospace and vehicle control. Multiplicative noise has also been observed in biological applications like sensorimotor control [20], [21], cell populations [22], [23], population migration [23], and immune systems [24]. Other applications include nuclear fission [23], power grids [25], communication channels [26], electrical networks [27], sampled data feedback [27], and climate models [28].

Second, the use of multiplicative noise induces robustness against parametric uncertainty. In [29], this is motivated by the lack of stability margins in output-feedback LQG [30]. This same fact was argued in a historical review by Doyle [31] to have motivated the development of $H_\infty$ methods. In fact, as we formalize later, there is a strong relation between the mean square stability (MSS) for multiplicative noise systems and robust stability [29], [32], [33].

Next, we provide a detailed overview of related work.

A. Related Work

1) Historical Overview: The history of multiplicative noise is extensive and has been investigated under various names. As linear dynamics with state-dependent noise, it was first

---

1558-2523 © 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
considered by [15], who developed an optimal control scheme. This led to several other early works, outlined in [19, p. 6]. Notably, the renowned discovery of the uncertainty threshold principle [34], thwarts system stabilization upon its violation. The term stochastic bilinear systems has also been used to describe multiplicative noise systems [23], [35], [27]. Most recent developments primarily involve control synthesis with indefinite stage costs [36], numerical solutions for generalized Riccati equations [19], output feedback [37], and application of $H_{\infty}$ methods to stochastic control [19], [38]. We similarly consider robustification against perturbations to the distribution of the disturbance (similar to [39]), but do so in a data-driven fashion.

2) Learning Control for Multiplicative Noise: In the top-down approach, several results from LQR have been generalized to multiplicative noise already. Policy iteration has been investigated in [16], [18], and [40] and policy gradient in [17] and [18].

Developments in the bottom-up approach remain limited due to challenges in identifying multiplicative noise dynamics. Except for partial identification in some cases [41], comprehensive system identification has only been investigated recently. In [42] and [43], a scheme based on deterministic linear system identification was developed by averaging over trajectories. Unlike our method, this method lacks the ability to incorporate prior model structure and has loose confidence bounds. The same authors also considered an identification scheme where the mean dynamics are identified first and the noise is characterized using bootstrapping [44]. Similarly, this scheme comes with few guarantees. Finally, Di and Lamperski [45] consider a least-squares estimator similar to ours. They consider scalar measurements, thwarting complete identification. Their confidence bounds are valid for single-trajectories (while we use multiple). We show experimentally that our method is also applicable in the single-trajectory case, but leave a formal proof for further work.

3) Mathematical Tools: Throughout this work, we illustrate what mathematical tools are applicable in the analysis of linear systems with multiplicative noise.

We express the underlying bilinear structure using tensor algebra [46], [47], facilitating the formulation of least squares estimators and studying their accuracy.

Additionally, we leverage completely positive (CP) operators, classically used to model quantum channels [33]. They were linked to multiplicative noise dynamics before in [19] and to the stability of linear systems and Lyapunov operators in [48], [49], and [50]. The relation between MSS and robust stability was developed originally for CP operators [33].

Our estimator is similar to those considered in compressed covariance sensing [51] and those in bilinear estimation [52].

To develop our concentration inequalities, we use modern tools from matrix concentration [53]. Our controller synthesis uses results from robust optimization [54]. Specifically, we use distributionally robust optimization [13] to interpolate between robust and stochastic control synthesis.

B. Contributions

The main contributions of this work are as follows.

1) We show how linear quadratic (LQ) control of multiplicative noise is equivalent to LQ control of CP dynamics.

2) We consider an identification procedure that can both include prior knowledge about the dynamics and also performs well in the model-free setting. The optimal estimate is accompanied by a tight confidence set.

3) We introduce a synthesis procedure for DR LQR.

4) We analyze the sample complexity of the scheme.

As mentioned above, CP operators have been applied to multiplicative noise before [49]. A formal equivalence of LQ control has not been shown, however. The other results all strongly rely on this equivalence.

The system identification problem solved in 2) was previously considered in [42] and [43]. However, their confidence bound is too conservative for practical purposes and there is no possibility to include prior information about how the uncertainty enters into the dynamics. Particularly when using longer trajectories in the data, their bound grows when more data is gathered, which is opposite to the behavior observed in experiments (cf., Fig. 5 and [43, p. 15]). Furthermore, unlike our bound in Theorem 5.4, their bound is not data-driven. It instead depends on unknown system constants. In [45], a partial ambiguity set is derived, but the full system dynamics are never identified. We consider 2) to be the most important contribution.

We also generalize our previous work to the model-free setting: 3) extends upon [14] and 4) extends upon [55], and applies the result to the DR control synthesis procedure.

Numerical experiments demonstrate practical applicability.

The rest of this article describes the construction, solution, and analysis of DR control for LQR under multiplicative noise and is organized as follows. We provide a problem statement in Section II and preliminary results in Section III. We then consider the nominal case for LQR under multiplicative noise, when the dynamics are known. Next, in Section V, we describe an ambiguity set based on state measurements; in Section VI-A, we solve the resulting DR problem; and in Section VI-B, we study convergence of the solution to the nominal controller as more data are gathered. We end the article with numerical results in Section VII. Finally, Section VIII concludes this article.

This article is accompanied by a technical report [56], including more detailed proofs, experiments, and a section on including a deterministic term in the prior model structure.

C. Notation Overview

Below, we provide an overview of the notation.

Let $\mathbb{N}$ denote the integers and $\mathbb{R}$ the (extended) reals. We introduce the shorthand $\mathbb{N}_{a,b} = \{a, \ldots, b\}$ for $a, b \in \mathbb{N}$, with $a \leq b$. For $x \in \mathbb{R}^m$, let $\|x\|_2 = \sqrt{x^T x}$ denote the Euclidean norm and $B_n \subset \mathbb{R}^n$ the unit-norm ball. Assuming column vectors, let $(x, y) = [x; y]$ denote the vertical concatenation between vectors.

For a random vector $w$ let $E_{\nu}[w]$ denote the expectation with respect to the measure $\nu$. The subscript $\nu$ is omitted when the true distribution is implied.

Matrices: For $Z \in \mathbb{R}^{m \times n}$ let $\lambda(Z) = (\lambda_1, \ldots, \lambda_m)$ (when $m = n$) denote the vector of eigenvalues in descending order. Let $\lambda_{\max}(Z) = \lambda_1$, $\lambda_{\min}(Z) = \lambda_n$, and $\rho(Z)$ be the spectral radius. Similarly, $\sigma(Z) = (\sigma_1, \ldots, \sigma_r)$ denotes the vector of nonzero singular values in descending order with $r = rk(Z)$ the rank. Denote by $Z^\dagger$ and $Z^*$ the pseudoinverse and the transpose.
respectively. Also, let \( \|Z\|_2 \) (\( \|Z\|_p \)) be the spectral (Frobenius) norm. Let \( \otimes \) denote the Kronecker product and \( \text{vec}(X) \in \mathbb{R}^{mn} \) the vectorization of \( X \) resulting from stacking the columns, and \( \text{unvec}(\cdot) \) its reverse. When \( m = n \) let \( \text{Tr}(Z) \) denote the trace.

For matrices \( X, Y \) of conformable dimensions, we use \( [X; Y] \) (\( [X]_Y \)) for vertical (horizontal) concatenation. Let \( \text{blockdiag}(X, Y) \) be the block diagonal concatenation of \( X \) and \( Y \) and let \( I_d \in \mathbb{R}^{d \times d} \) be the identity. Elements of \( M \in \mathbb{R}^{m \times n} \) (and vectors \( x \in \mathbb{R}^d \)) are indexed using \( [X]_{ij} \) (\( [x]_i \)) for \( i \in [1:m], j \in [1:n] \). We can select columns (rows) using colon notation \( [X]_{:, j} \) (\( [X]_{i,:} \)). Moreover, to take the first \( d \) rows (columns) we write \( [X]_{d,:} \) (\( [X]_{:,d} \)).

We denote by \( S^d \) the set of symmetric \( d \times d \) matrices and by \( S^d_{++} \) (\( S^d_{+} \)) the positive (semi)definite matrices. For \( P, Q \in S^d \), we write \( P \succeq Q \) (\( P \succeq Q \)) to signify \( P - Q \in S^d_{++} \) (\( P - Q \in S^d_{+} \)). Let \( \delta_d = (d + 1)d/2 \) denote the degrees of freedom of \( S^d_{++} \) and \( \text{vec}_d(X) \in \mathbb{R}^{d^2} \) denotes the symmetrized vectorization (cf., Definition 3.1) such that \( \text{vec}_d(X) \) \( \text{vec}_d(X) = \|X\|_F^2 \), with \( \text{unvec}(\cdot) \) its inverse. Let \( \otimes_s \) the symmetrized Kronecker product (cf., Definition 3.2).

**Tensors**: We use bold calligraphic letters to denote (third-order) tensors \( \mathcal{A} \in \mathbb{R}^{p \times q \times r} \). We follow the notation in [46] and provide stand-alone definitions in Section III-B. A tensor is indexed as \( [\mathcal{A}]_{ijk} \) with \( i \in [1:p], j \in [1:q], k \in [1:r] \). We generalize the colon notation of taking rows and columns for matrices for selecting subtensors (e.g., \( [\mathcal{A}]_{i,\cdot,\cdot} \)) and use \( (\otimes_s) \times_n \) to denote the \( n \)-mode (vector) product and \( \mathcal{A}_{(n)} \) for the \( n \)-mode matricization. Let \( \mathcal{A} = \{X_1, X_2, X_3\} := \{X_{11}X_1 \times X_2 \times X_3 \} \) denote the Tucker operator (cf., Remark 3.10).

Multiplicative noise: For convenience we illustrate the use of tensors for autonomous multiplicative noise dynamics given as \( x_{t+1} = (\sum_{i=1}^{n_w} A_i[w_i])x_t, \) with \( x_t \in \mathbb{R}^{n_x} \) and \( w_t \in \mathbb{R}^{n_w} \). The associated tensors is \( \mathcal{A} \in \mathbb{R}^{n_x \times n_x \times n_w} \) populated as \( [\mathcal{A}]_{i,:} = A_i \) for \( i \in [n_w] \). The matricizations are

\[
\mathcal{A}(1) = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix}, \quad \mathcal{A}(2) = \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix}^T,
\]

\[\mathcal{A}(3) = \begin{bmatrix} \text{vec}(A_1) & \text{vec}(A_2) & \text{vec}(A_3) \end{bmatrix}^T.
\]

Using this notation \( x_{t+1} = [\mathcal{A} ; I_n \times x_t, w_t ] = \mathcal{A}(1)(w_t \otimes x_t) = x_t \mathcal{A}(2)(w_t \otimes I_n) = w_t \mathcal{A}(3)(x_t \otimes I_n) \).

**II. PROBLEM STATEMENT**

In this work, we consider linear systems with input- and state-multiplicative noise given for all \( t \in \mathbb{N} \) as

\[
x_{t+1} = A(v_t)x_t + B(v_t)u_t \tag{1}
\]

where \( x_t \in \mathbb{R}^{n_x} \) is the state, \( u_t \in \mathbb{R}^{n_u} \) the input and \( v_t \in \mathbb{R}^{n_v} \) the disturbance, which follows an i.i.d. random process. The matrices are given by \( A(v) := \sum_{u=1}^{n_u} [v]A_u \) and \( B(v) := \sum_{v=1}^{n_v} [v]B_v \).

The primary goal is to study solutions of the following stochastic LQR problem:

\[
\text{minimize } \mathbb{E} \left[ \sum_{t=0}^{\infty} x_t^T Q x_t + u_t^T R u_t \right], \tag{LQR*}
\]

subject to (1), for given \( x_0 \) and \( Q \succ 0 \) and \( R \succ 0 \).\(^1\)

\(^1\)These assumptions are not necessary for solvability [36], yet are added here for convenience.

In practical applications, the dynamics (1) are often unknown. Hence, we introduce a system identification scheme to estimate (1), supporting both the setting where the modes \( A_i, B_i \) are unknown and the setting where they are (partially) known. This information will be encoded through the use of a model tensor \( \mathcal{M} \) described in Section IV-A.

Besides the modes, characterizing the distribution of \( v_t \) is a more considerable challenge. We, however, are aided by the following fact. By linearity, the cost of \( (LQR_*) \) only involves the second moments of the states \( X_t = E[x_t x_t^T] \) and the inputs \( U_t = E[u_t u_t^T] \). In this work, we thus tackle LQR in terms of the moment dynamics

\[
X_{t+1} = E(Z_t), \quad \text{with } Z_t = E[z_t z_t^T] \tag{2}
\]

for augmented state \( z_t = (x_t, u_t) \in \mathbb{R}^{n_x+n_u} \). These dynamics are easy to estimate, compared to (1), yet they contain all the information required to solve the stochastic LQR problem. We view (2) as an implicit definition for the operator \( E \)—the explicit definition being (1)—and argue in Section IV-A how the fundamental operator \( E \) (i) is linear in \( z_t \), characterizing its matrix \( E \) in terms of the quantities (1), (ii) is completely positive (CP) (cf., Definition 3.4 and Lemma 3.12), (iii) has a bilinear structure since \( E \) depends linearly on \( V = E[\begin{pmatrix} v_t \end{pmatrix} \begin{pmatrix} v_t \end{pmatrix}] \). These properties serve as the central theme of this work.

Since we will never get a completely accurate estimate of the matrices, it makes sense in safety critical scenarios to use data to construct a set of likely distributions \( \mathcal{S} \) containing the true distribution with high probability (see the formal setup in [57]). Given \( \mathcal{S} \), we then solve the DR LQR problem

\[
\text{minimize } \mathbb{E}_\mathcal{S} \left[ \sum_{t=0}^{\infty} x_t^T Q x_t + u_t^T R u_t \right], \tag{LQR*}
\]

which upper bounds the cost of \( (LQR_*) \) with high probability. In this work, only the second moment \( V = E[\begin{pmatrix} v_t \end{pmatrix} \begin{pmatrix} v_t \end{pmatrix}] \) affects the problem. So we formulate the ambiguity set over moments.

**III. PRELIMINARY RESULTS**

We introduce the main tools used throughout our work here. Specifically, this section presents operators over the space of symmetric matrices in Section III-A and tensor algebra in Section III-B.

**A. Operators Over Symmetric Matrices**

As argued later, the second moment of the disturbance \( v \) is identified by solving a linear system of equations, where the variable of interest is a symmetric matrix.

Useful tools when dealing with such problems are the Kronecker product \( \otimes \) [58, Def. 4.2.1], the vector operator \( \text{vec}(\cdot) \) [58, Def. 4.2.9] and its inverse \( \text{unvec}(\cdot) \). These are connected through the fundamental property [58, Lemma 4.3.1]

\[
\text{vec}(X^TV^T) = (V^T \otimes V)\text{vec}(X),
\]

for any \( U \in \mathbb{R}^{m \times n}, V \in \mathbb{R}^{k \times k} \), and \( X \in \mathbb{R}^{n \times k} \).

This property is often used to rewrite linear equations over matrices in terms of vectors. When \( X \in \mathbb{R}^d \), however, we often run into noninvertible systems since \( \text{vec}(X) \in \mathbb{R}^{d^2} \) has unnecessary degrees of freedom. After all, a symmetric matrix is uniquely determined by \( d = (d+1)/2 \) values.
To exploit this fact, we will use [59, Def. E.1].

**Definition 3.1:** For any \( X \in S^d \), define \( \text{vec}_a : S^d \rightarrow \mathbb{R}^{d^3} \)
\[
\text{vec}_a(X) := ([X]_{11}, \sqrt{2}[X]_{21}, \ldots, \sqrt{2}[X]_{d1})
\]
and \( Q_d \in \mathbb{R}^{d \times d} \) be such that \( \text{vec}_a(X) = Q_d \text{vec}(X) \).

It is easy to verify that \( Q_d \) is a unique matrix for which \( Q_d Q_d^T = I_d \) [59, §E.2]. We can then introduce the symmetrized Kronecker product [59, Def. E.3].

**Definition 3.2:** For any \( U, V \in \mathbb{R}^{m \times n} \), let
\[
V \otimes_n U = \frac{1}{2} Q_m (U \otimes V + V \otimes U) Q_m^T
\]
By definition, we have a generalization of (3) [59, Def. E.3]
\[
\text{vec}_a(ZX Z^T) = (Z \otimes Z) \text{vec}_a(X)
\]
We will also consider more general linear operators on symmetric matrices, \( S : S^n \rightarrow S^m \), which can be expressed for any \( X \in S^n \) as
\[
S(X) = \text{vec}_a(S \text{vec}_a(X))
\]
for a unique \( S \in \mathbb{R}^{m \times n^2} \). The set of such operators is denoted \( SL^{m \times n} \). The adjoint \( S^* \), defined to satisfy \( \text{Tr}(S^*(Y)X) = \text{Tr}(Y^T S(X)) \), has matrix \( X^T S \). Similar to [55], we will often need to bound the spectral norm of the image of such matrix operators. As such, we introduce the following norm on \( SL^{m \times n} \):
\[
\| S \|_2 := \max_{X \in S^n} \{ \| S(X) \|_2 : \| X \|_2 \leq 1 \}
\]
which for which we have \( \| S(X) \|_2 \leq \| S \|_2 \| X \|_2 \) for any \( X \in S^n \). This norm is bounded in terms of the spectral norm of \( S \):

**Lemma 3.3:** Let \( S \in SL^{m \times n} \) with matrix \( S \). Then
\[
\| S \|_2 \leq \sqrt{m} \| S \|_2 \tilde{s}
\]
**Proof:** Deferred to [56, Appendix A].

We are particularly interested in linear operators that preserve the positive semi-definiteness of the argument. That is, a map \( S : S^n \rightarrow S^m \) is positive iff \( S(X) \geq 0 \) for all \( X \geq 0 \). More specifically, we consider a specific type of positive map:

**Definition 3.4:** A \( S \in SL^{m \times n} \) is completely positive (CP) [60] if, for some \( A_i \in \mathbb{R}^{m \times n} \) with \( i \in \mathbb{N}_1 : r \)
\[
S(X) = \sum_{i=1}^r A_i X A_i^T.
\]

The matrix of \( S \) is \( S = \sum_{i=1}^r (A_i \otimes_n A_i) \). We write \( S \in CP^{m \times n} \). We specialize Lemma 3.3 for CP operators [48, eq. (2.1)]:

**Lemma 3.5:** Let \( S \in CP^{m \times n} \). Then, \( \| S \|_2 = \| S(I) \|_2 \).

Like for linear dynamics, stability of CP operators can be defined in terms of the spectral radius.

**Lemma 3.6:** For any \( S \in CP^{m \times n} \) the following assertions are equivalent:

a) \( \rho(S) := \rho(S) < 1 \);

b) \( \| S^t(X) \|_2 \rightarrow 0 \) as \( t \rightarrow \infty \) \( \forall X \succeq 0 \).

**Proof:** The proof is analogous to [35, Prop. 1].

The spectral radius of a CP map \( S \) is related to the outer spectral radius \( \rho(A_1, \ldots, A_r) := \rho(S(\sum_{i=1}^r A_i \otimes A_i)) \) for square matrices \( \{A_i\}_{i=1}^r \). This radius bounds the joint spectral radius
\[
\rho(A_1, \ldots, A_r) := \lim_{t \rightarrow \infty} \left\{ \sup_{1 \leq i_1, \ldots, i_t \leq r} \| A_{i_1} \ldots A_{i_t} \|_2^{1/t} \right\}.
\]

Specifically [33, Cor. 1.5]:

**Theorem 3.7:** Consider the matrices \( \{A_i\}_{i=1}^r \). Then
\[
\hat{\rho}(A_1, \ldots, A_r) := \rho(A_1) / \sqrt{r} \leq \rho(A_1, \ldots, A_r) \leq \hat{\rho}(A_1, \ldots, A_r).
\]

Note that the joint spectral radius describes the stability the dynamics of a switched linear system. That is, dynamics \( \tau + i \tau_0 \) are stable for any sequence of matrices \( A_i \) from \( \{A_i\}_{i=1}^r \) iff \( p < 0 \) [62, Cor. 1.1]. Therefore, Theorem 3.7 gives a strong relation between the stability of such systems and the stability of \( S \) (and, as we will see later, the mean square stability of linear dynamics with multiplicative noise).

We finally consider a Lyapunov criterion for stability.

**Proposition 3.8:** Let \( S \in CP^{d \times d} \). Then
i) If \( \rho(S) < 1 \), then \( \forall H > 0 \)
\[
\exists P \succeq 0 : P - S^* P = H.
\]
Moreover, \( P \) is unique, positive definite, and \( \text{Tr}(PX) = \text{Tr}(H \sum_{i=1}^r S^t(X)) \) \( \forall X \succeq 0 \).

ii) If (6) for some \( H > 0 \), then \( \rho(S) < 1 \).

**Proof:** For linear systems with multiplicative noise, this statement was shown in [35, Lemma 1-2]. The proof for CP operators is similar and is included in [56].

**B. Tensor Algebra**

The dynamics in (1) inherently have a bilinear structure. This warrants the use of multilinear structures and specifically third order tensors. In this section, we introduce tensors and show how they are used to model CP operators.

Let \( T \in \mathbb{R}^{q_1 \times q_2 \times q_3} \) denote a generic third-order tensor. A good introduction is given in [46, 63, §12.4]. We restate some results here for completeness.

We index tensors using \( [T]_{ijk} \in \mathbb{R} \). Subtensors are selected with \( \text{colon notation} \) [63, §12.4] (e.g., \( [T]_{ijk} \in \mathbb{R}^{q_1} \)). We take a slice along mode- \( n \) for \( n \in \{1, 2, 3\} \) using \( [T]_{i \ast k} \in \mathbb{R}^{q_n \times q_{123-n}} \). Consider \( n \)-mode matricizations \( T_{(n)} \in \mathbb{R}^{q_n \times (q_{123-n} q_{123-n})} \)
\[
T_{(n)} := [\text{vec}([T]_{i \ast k})^T], \ldots, \text{vec}([T]_{i \ast k})^{(q_{123-n})}.
\]

These matricizations extract the linearity of the multilinear tensor operator in one of its arguments. This is illuminated by introducing the mode- \( n \)-product. This operation multiplies \( T \in \mathbb{R}^{q_1 \times q_2 \times q_3} \) along axis \( n \) with matrix \( X \in \mathbb{R}^{p_n \times q_n} \) resulting in a tensor where \( q_n \) is repeated by \( p_n \). The definition can be given in terms of matricizations
\[
(T \times_n X)_{(n)} := X^T T_{(n)}.
\]

Since matricization is one-to-one, this uniquely defines the tensor \( T \times_n X \). Alternatively, e.g., for \( n = 2 \), we have \( [T \times_2 X]_{ijk} = \sum_{l=1}^{q_3} [X]_{ijl} [T]_{l \ast k} \) with \( i \in \mathbb{N}_{1:q_1} \), \( j \in \mathbb{N}_{1:q_2} \), and \( k \in \mathbb{N}_{1:q_3} \). The extraction of linearity in one of the arguments is most explicit in (8). Similarly, the mode- \( n \)-vector product is \( T \times_n x = \sum_{j=1}^{q_3} [x]_j [T]_{i \ast k} \). So the order of the tensor is reduced to \( (T \times_n x) \in \mathbb{R}^{q_n \times q_{123-n}} \).

We can extend the single mode- \( n \)-product to the multilinear one, similarly to (8).

**Definition 3.9 (Tucker product):** For \( n \in \{1, 2, 3\} \), let \( X_n \in \mathbb{R}^{p_n \times q_n} \). We call \( [T; X_1, X_2, X_3] \) the Tucker product with
\[
\mathbf{Y} := [(T \times_3 X_3) \times_2 X_2] \times_1 X_1
\]
\[
[(T; X_1, X_2, X_3) \in \mathbb{R}^{p_1 \times p_2 \times p_3}.
\]
When $p_n = 1$, which is the case when $X_n = x_n^r$ for (column) vector $x_n \in \mathbb{R}^{p \times n}$ then that axis is removed, reducing the order of $\mathcal{Y}$. That is, $x_n$ is replaced by $x_n^r$. As is common in literature, we omit the transposes in this case, writing $[T; x_1, x_2, x_3]$. 

Remark 3.10: The order of operations in (9) is important when acting on vectors. After all, since $\mathcal{T} \times x_2 x_3 \in \mathbb{R}^{q_3 \times q_3}$ writing $(\mathcal{T} \times x_2 x_3) \times x_3$ is not defined. Instead, we would need to write $(\mathcal{T} \times x_2 x_3) \times x_3 = (\mathcal{T} \times x_2) x_3$. The Tucker operator helps to avoid such complexities, since we assume the order is reduced after all mode-$n$ products have been completed.

We frequently use a generalization of (8) [47, Prop. 3.7].

Proposition 3.11: For $\mathcal{Y} \in \mathbb{R}^{p_1 \times p_2 \times p_3}$ as in Definition 3.9

$$\mathcal{Y} = X_n \mathcal{T}(n)(X_j \times X_i)$$

for $n, i, j \in \{1, 2, 3\}$ with $j > i$ and $n \neq i \neq j$.

A graphical illustration of (10) is given in Fig. 1. Note, how the values along the third axis are split up both by the matricization and the Kronecker product.

We can describe CP maps using tensors.

Lemma 3.12: Any $A \in \mathbb{C}^{p \times n \times r}$ and $W \in S_+ ^n$. We will refer to $A$ as the model tensor and $W$ as the parameter matrix.

Proof: Deferred to [56, Appendix A].

Remark 3.13: The dependence of $S$ on $W$ in (11) can be made explicit by writing $S(W; X)$. That is $S(W; \cdot) : \mathbb{R}^{m \times n} \to \mathbb{R}^r$. We distinguish the parameters of the map from its arguments using a semicolon. So $S(W; P)$ denotes the adjoint w.r.t. the argument, not the parameter $W$.

We show that $S(W; X)$ is also CP in its parameter $W$.

Corollary 3.14: Let $S(W; X)$ as in (11). Then

i) $\forall X \succeq 0$, $S(\cdot; X) \in \mathbb{C}^{p \times n \times r}$

ii) $S(W; X) \succeq S(W; X)$ if $W \succeq W$.

Proof: Deferred to [56, Appendix A].

We will see later, in the setting of multiplicative noise dynamics, that Corollary 3.14 is useful for finding sufficient conditions for stability. Specifically, it allows constructing dynamics whose second moments dominate those of the true dynamics in the psd.

There is a bilinear structure present in (11). This becomes clear when considering $\text{Tr}[PS(W; X)]$, which is linear in $P$, $X$ and $W$. Hence, there is a version of Proposition 3.11 for CP maps (cf., [56, Prop. A.1]), which allows different characterizations of $S$ and $\mathcal{S}$. 

We define Kronecker products on $\mathcal{T} \in \mathbb{R}^{q_1 \times q_2 \times q_3}$ [64]

$$[\mathcal{T} \otimes \mathcal{T}]_{(i+j-i)q_3} = [\mathcal{V}]_{i \cdot} \otimes [\mathcal{V}]_{j \cdot}$$

for $i, j \in \{1, 2, 3\}$.

Also, let $\mathcal{V} \otimes \mathcal{V} = [\mathcal{V} \otimes \mathcal{V}^T]_{q_1 \times q_2 \times q_3}$.

Next, we state a generalization of a property for the Kronecker product between matrices, which is used to rewrite CP operators.

Lemma 3.15: Suppose $\mathcal{T} \in \mathbb{R}^{q_1 \times q_2 \times q_3}$ and $X_i \in \mathbb{R}^{p \times q_i}$ for $i \in \{1, 2, 3\}$ and let $\mathcal{Y} = [\mathcal{T}; X_1, X_2, X_3]$. Then

$$\mathcal{Y} \otimes \mathcal{Y} = [\mathcal{T} \otimes \mathcal{T}; X_1 \otimes X_1, X_2 \otimes X_2, X_3 \otimes X_3]$$

Proof: Deferred to [56, Appendix A].

Corollary 3.16: Let $S(W; X)$ be CP and parameterized as in (11) with some $\mathcal{Y}$. Then $S(W; X) = \mathcal{V}_{(2)}(W \otimes X) \mathcal{V}_{(2)}$. Moreover its matrix $S = (\mathcal{V} \otimes \mathcal{V}) \times_3 \text{vec}(W)$.

Proof: Deferred to [56, Appendix A].

IV. KNOWN DISTRIBUTION

We consider the nominal setting for multiplicative noise, where the distribution is known. The (CP) operator $\mathcal{E}$ from (2), which fully describes the second-moment dynamics of (1), is examined. We prove that it is (i) linear in $Z_i$; (ii) completely positive; and (iii) is linear in $E[v_t, v_t^T]$. Moreover, we argue that identifying $\mathcal{E}$ is sufficient to describe both stability and quadratically optimal control.

A. Model Tensor

We first illustrate how the previous tools interface with multiplicative noise. Consider the true model tensor $\mathcal{V} \in \mathbb{R}^{n_x \times n_x \times n_x}$, which is populated as $[\mathcal{V}]_{i \cdot} = [A_i, B_i]$ for $i \in \mathbb{N}_{1:n_v}$, using the modes in (1). This splits up the dynamics into two unknowns: the tensor $\mathcal{V}$ and the distribution of $v_t$.

To encode known information about $\mathcal{V}$, we introduce the model tensor $\mathcal{M} \in \mathbb{R}^{n_x \times n_x \times n_w}$ and the auxiliary random vector $w_t \in \mathbb{R}^{n_w}$. When the modes are known, we can simply pick $\mathcal{M} = \mathcal{V}$ and $w_t = v_t$. However, when $\mathcal{V}$ is only partially known, we can still select a $\mathcal{M}$ that has sufficient modeling capacity. Specifically, we would like that for any realization of the disturbance $v_t$ and hence of the random matrices $A_t(v_t)$ and $B_t(v_t)$, there exists a realization of $w_t$ producing the same matrices by the modes encoded in $\mathcal{M}$.

A sufficient assumption for this property is as follows.

Assumption 1 (Model equivalence): Consider $\mathcal{V} \in \mathbb{R}^{n_x \times n_x \times n_w}$ with $n_x := n_x + n_v$ s.t. $[\mathcal{V}]_{i \cdot} = [A_i, B_i]$ for $i \in \mathbb{N}_{1:n_v}$ and $[A_i, B_i]_{n+1} = n_v$ as in (1). Assume $\mathcal{M} \in \mathbb{R}^{n_x \times n_x \times n_x}$ with $n_w \in \mathbb{N}$ is such that

$$rk(\mathcal{M}_{(3)}) = rk([\mathcal{M}_{(3)}; \mathcal{V}_{(3)}]) = n_w.$$ 

Here, $\mathcal{M}_{(3)} \in \mathbb{R}^{n_u \times n_u \times n_z}$ and $\mathcal{V}_{(3)} \in \mathbb{R}^{n_u \times n_u \times n_z}$ denote the 3-mode matricizations as in (7).

Note that $\mathcal{M}$ can be partitioned into $\mathcal{M} := [\mathcal{M}]_{1,1,1}$ and $\mathcal{B} := [\mathcal{M}]_{n_x+1,1,1}$. These parts serve as an analogy of $A$ and $B$ for linear time-invariant systems. When $\mathcal{M} = \mathcal{V}$, we have $[A_i, B_i]_{n+1} = n_v$. We can still make this partitioning, we can show that

$$x_t+1 = [\mathcal{M}; I_{n_x, z_t}, w] + [\mathcal{B}; I_{n_x, v_t}, w]$$

(13) describes the same dynamics as (1), given a suitable choice of $w_t$. We formalize equivalence of (13) and (11) below.

Lemma 4.1: When Assumption 1 holds and $w_t = (\mathcal{M}_{(3)})^T \mathcal{V}_{(3)} v_t \forall t \in \mathbb{N}$, then (13) produces the same trajectories as (1).

Proof: The statement in Assumption 1 implies that the columns of $\mathcal{V}_{(3)}$ are in the range of $\mathcal{M}_{(3)}$. This implies that $\mathcal{M}_{(3)} w_t = \mathcal{M}_{(3)}^T (\mathcal{M}_{(3)})^T \mathcal{V}_{(3)} v_t = \mathcal{V}_{(3)} v_t$, since
$\mathcal{M}^\top_{(3)}(\mathcal{M}_{(3)}^\top)^\dagger$ is an orthogonal projection onto the range of $\mathcal{M}_{(3)}^\top$.

Using (8) on both sides of $w_3^T\mathcal{M}_{(3)} = v^T_1\mathcal{V}_{(3)}$ and reshaping gives $(\mathcal{M} \times_3 w_3) = (\mathcal{V} \times_3 v_1^T)$. Replacing $\times_3$ with $\times_3$ only alters the dimensions, without changing the total number of elements. So $(\mathcal{M} \times_3 w_3) = (\mathcal{V} \times_3 v_1)$. Then, by the definitions of $\times_3$, $\mathcal{V}$ and $A(v)$, $B(v)$ we have, respectively,

$$\mathcal{V} \times_3 v = \sum_{i=1}^{n_v} [v]_i [\mathcal{V}]_{:,i} = \sum_{i=1}^{n_v} [v]_i [A_i, B_i] = [A(v), B(v)].$$

Finally, by Definition 3.9, $[\mathcal{M}; I_{n_x}, z_t, w_t] = (\mathcal{M} \times_3 w_3) \times_2 [A(v), B(v)]z_t = A(v_t)x_t + B(v_t)u_t$. So the produced trajectories will be the same. \hfill $\square$

**Example 4.2:** We illustrate the need for Assumption 1. Assume $n_x = 2$, $n_v = 1$, and $n_\tau = 3$ with true modes

$$A_1 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

and $B_1 = B_2 = 0$. In this case, $\mathcal{V}_{(3)}$ is given as

$$\begin{bmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{bmatrix} = \begin{bmatrix} \text{vec}(A_1) \end{bmatrix}^\top \begin{bmatrix} \text{vec}(B_1) \end{bmatrix}^\top \begin{bmatrix} \text{vec}(A_2) \end{bmatrix} \begin{bmatrix} \text{vec}(B_2) \end{bmatrix} \begin{bmatrix} \text{vec}(A_3) \end{bmatrix} \begin{bmatrix} \text{vec}(B_3) \end{bmatrix}. $$

Say a designer, through some erroneous prior analysis, assumed that the top-left component in $A_1$ was equal to zero instead of two. This corresponds to the following model tensor:

$$\mathcal{M}_{(3)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 2 \end{bmatrix}$$

for this choice, Assumption 1 fails and so will Lemma 4.1. To see this, take $v_t = (1, 0, 0)$. Then vec$([A(v), B(v)]) = (2, 1, 0, 0, 0, 0)$. Observe that such a vector cannot be constructed by any linear combination of the rows of $\mathcal{M}_{(3)}$. So there is no distribution for $w_t$ that reproduces the true dynamics.

Informally, Assumption 1 requires that the span of $\mathcal{M}_{(3)}$ should be sufficiently large to model the outputs of $A(v)$ and $B(v)$ given by realizations of $v \in \mathbb{R}^n_v$. Note that this is always possible, by taking $n_w = n_x(n_x + n_\tau)$ and $\mathcal{M}_{(3)} = I_{n_w}$. We refer to this as the model-free basis, since no prior structural assumptions are imposed on the dynamics. If we know the true modes $A_t$ and $B_t$, we can simply select $\mathcal{M} = \mathcal{V}$. As such, $\mathcal{M}$ is used to introduce prior information into the dynamics.

Since the sequence $v_t$ is i.i.d., the sequence $w_t$ as defined in Lemma 4.1 is i.i.d. too. We thus omit it as the distribution remains constant. We can then use the transformation in Lemma 4.1 to find the true $W = \mathbb{E}[w_1 w_3^\top]$ in terms of $V = \mathbb{E}[v_1 v_3^\top].$

$$W = (\mathcal{M}_{(3)}^\top)^\dagger \mathcal{V}_{(3)}^\top \mathcal{V}_{(3)} \mathcal{M}_{(3)}^\top.$$  (14)

This allows translating the ground truth dynamics to another model tensor and will aid in the validation of our method during the experiments in Section VII.

We next illustrate the connection with CP operators by considering the dynamics of $X_t = \mathbb{E}[x_t x_t^\top]$ as in (2). By Proposition 3.11, $[\mathcal{M}; I, z_t, w_t] = \mathcal{M}_{(1)}^\top(w_t \otimes z_t)$. Taking the outer product, using $(w_t \otimes z_t)(w_t \otimes z_t^\top) = (w_tw_t^\top \otimes z_t z_t^\top)$ and taking the expectation allows us to derive

$$X_{t+1} = \mathcal{E}(W; Z_t) := \mathcal{M}_{(1)}^\top(W \otimes Z_t) \mathcal{M}_{(1)}^\top$$  (15)

where $X_t = \mathbb{E}[x_t x_t^\top]$, $Z_t = \mathbb{E}[z_t z_t^\top]$, and $W = \mathbb{E}[w_t w_t^\top]$. We write $\mathcal{E}()$ when the specific value of $W$ is unimportant. The operator $\mathcal{E} : \mathbb{S}^{n_x^2} \rightarrow \mathbb{S}^{n_x^2}$ is clearly CP by Lemma 3.12. It is linear with its matrix characterized by Corollary 3.16. Moreover it has a bilinear structure, depending linearly on $W$ (and by (14) on $V$). So we confirmed the claims from the problem statement. We next discuss nominal stability and LQR, before introducing the system identification scheme.

**B. Stability**

We consider the stability of the autonomous case of (13)

$$x_{t+1} = [\mathcal{A}; I_{n_x}, x_t, u_t].$$  (16)

Stability of stochastic systems is intimately related to the convergence of random variables, which can be defined in several ways [65, §7.2]. We specifically consider the convergence to zero of the second moment [65, Def. 1.a].

**Definition 4.3:** Consider the autonomous system (16). It is considered mean-square stable (MSS) if

$$\mathbb{E}[x_t x_t^\top] \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty.$$  

According to the definition, MSS is equivalent to the stability of

$$X_{t+1} = \mathcal{F}(X_t) := \mathcal{A}_{(1)}(W \otimes X_t) \mathcal{A}_{(1)}^\top$$

where $X_t = \mathbb{E}[x_t x_t^\top]$ as before and $\mathcal{F}$ is the autonomous version of $\mathcal{E}$. Clearly this system is stable for all $X_0 \geq 0$ iff $\rho(\mathcal{F}) < 1$ by Lemma 3.6. Therefore:

**Corollary 4.4:** The system (16) is MSS iff $\rho(\mathcal{F}) < 1$.

We can verify stability with the Lyapunov condition Proposition 3.8 and, as mentioned earlier, the stability of (16) is related to robust stability of the switching system with modes $\{A_i\}_{i=1}^{n_\tau}$ in the sense of Theorem 3.7. This relation was observed before in [29] and [32], yet not fully characterized.

**C. Linear Quadratic Regulation**

The goal of this section is twofold (i) we argue how LQ control of CP dynamics is equivalent to $(\mathcal{LQR}_*)$; and (ii) we show how the optimal policy is determined by solving a SDP, generalizing the result of [66].

We start by stating the CP equivalent of $(\mathcal{LQR}_*)$.

$$\begin{array}{ll}
\text{minimize} & \sum_{t=0}^{\infty} \text{Tr}[Z_t H] \\
\text{subject to} & Z_t = \begin{bmatrix} X_t & V_t^\top \\ V_t & U_t \end{bmatrix} \succeq 0 \quad (\mathcal{LQR}_*_{cp}) \\
& X_{t+1} = \mathcal{E}(Z_t) \quad \forall t \in \mathbb{N},
\end{array}$$

with $H := \text{blkdiag}(Q, R) > 0$. The psd. constraint on $Z_t$ ensures that it acts like the second moment of a random vector.

We show that $(\mathcal{LQR}_*)$ is equivalent to $(\mathcal{LQR}_*_{cp})$

**Proposition 4.5:** Let $\text{Val}(\cdot)$ denote the optimal value. Then

$$\text{Val}(\mathcal{LQR}_*) = \text{Val}(\mathcal{LQR}_*_{cp})$$

given Assumption 1. Moreover, sequence $(Z_t)_{t \in \mathbb{N}}$ is feasible for $(\mathcal{LQR}_*_{cp})$ iff there is a feasible sequence in $(\mathcal{LQR}_*) u_t = \cdots$
\(K \neq \delta_t \text{ s.t. } Z_t = E[z_t z_t^\top] \) where \(z_t = (x_t, u_t)\). Here, \(K \in \mathbb{R}^{n_x \times n_x}\) and \(\delta_t \in \mathbb{R}^{n_u}\) is a random vector \(\forall t \in \mathbb{N}\).

Proof: Defered to Appendix A.

To find the exact solution of \((\mathcal{LQQR}_{cp})\), similarly to the usual LQR, we need to consider a Riccati equation

\[
R(P) := Q + F^*(P) - (H^*(P))^{-1} R + G^*(P)^{-1} H^*(P)
\]

where

\[
F^*(P) = A_{(2)}(W \otimes P)A_{(2)}, \quad G^*(P) = B_{(2)}(W \otimes P)B_{(2)}, \quad H^*(P) = B_{(2)}(W \otimes P)A_{(2)}
\]

with tensors \(A\) and \(B\) as in (13). We use adjoin CT operators to be consistent with the 2-mode matricizations as they occur in Corollary 3.16.

Remark 4.6: The operators \(F^*, H^*, G^*\) form the blocks of \(E^*\) when it is partitioned similarly to \(Z_t\) (cf., Lemma 1.1).

In the deterministic setting with \(n_{w_t} = 1\) and \(u_{t_1} = 1\) (so \(E[W] = 1\)), we have \(F^*(P) = A_1^* PA_1, H^*(P) = B_1^* PA_1\) and \(G^*(P) = B_1^* PB_1\). Hence, we recover the classical Riccati operator as a special case.

We introduce the closed-loop operators as

\[
\Pi_K(X) := [I \cdot K]X[I \cdot K^*], \quad \mathcal{E}_K(X) := \mathcal{E}([\Pi_K(X)]).
\]

The policy \(Z_t = \Pi_K(X_t)\) is recovered when \(u_t = K x_t\) in the equivalent \((\mathcal{LQQR}_s)\) (cf., Theorem 4.5).

We can then state the following theorem.

Theorem 4.7: Assume there is some \(K\) such that \(\rho(\mathcal{E}_K) < 1\) and that \(H := \text{blkdiag}(Q, R) \succ 0\). Then \(Z_t = \Pi_K(X_t)\) is an optimal policy, with \(K_t = \begin{bmatrix} (R + G^*(P))^{-1} H^*(P) \\ P_s \end{bmatrix}\) and \(P_s \succ 0\) solves the Riccati equation \(P_s = \mathcal{R}(P_s)\). Moreover

i) \(\rho(\mathcal{E}_K) < 1\);

ii) \(\mathbf{Val}(\mathcal{LQQR}_{cp}) = \text{Tr}[P_s X_0]\).

Proof: This result was shown for multiplicative noise in [14] and [67] and generalized to the CP setting by Theorem 4.5. Nonetheless, a full proof is included in the technical report [56].

Remark 4.8: Here, the existence of a \(K\) such that \(\rho(\mathcal{E}_K) < 1\) is referred to as stability. Among other things, it implies the system being below the uncertainty threshold [34].

Even though we directly tackled second-moment dynamics in Theorem 4.7, the optimal policy \(Z_t = \Pi_K(X_t)\) is realizable for multiplicative noise dynamics using \(u_t = K_t x_t\). So this controller is also optimal for the true dynamics (i.e., Assumption 1) and generalized to the CP setting by Theorem 3.7.

The final question is how to find a \(P_s\) satisfying (17). Numerical solution of such equations is considered in detail in [49]. We consider a reformulation as an SDP, akin to [66].

Theorem 4.9: For \(F^*, H^*, G^*\) as in (17), consider

maximize \(\text{Tr}[PX_0]\)

subject to

\[
\begin{bmatrix} P - F^*(P) & -(H^*(P))^{-1}R \\ -H^*(P) & \sigma P \end{bmatrix} \preceq \begin{bmatrix} Q & R \\ R & 0 \end{bmatrix}
\]

with \(Q \succ 0\) and \(R \succ 0\). Then

i) (19) is bounded iff \(\exists K : \rho(\mathcal{E}_K) < 1\);

ii) For bounded (19), the \(P_s\) that solves (17) is an optimizer.

Moreover, for \(X_0 \succ 0, P_s\) is the unique solution.

Proof: The proof is analogous to the one in [66] and [14, Prop. 3]. A full proof in the CP case is deferred to [56].

So the SDP in Theorem 4.9 both provides a way of solving the Riccati equation and a way to verify stabilizability.

V. IDENTIFYING THE SECOND MOMENT

Our goal in this work is to generalize Theorem 4.7 to the data-driven DR setting \((\mathcal{LQR})\), while retaining stability for a linear controller. This problem was considered in [14], for sub-Gaussian normalized disturbance \(\xi (i.e., \xi := \text{Var}[w]^{-1/2}(w - E[w]))\). Therefore, Coppens et al [14] supported, among others, Gaussian \(w\) or bounded \(\xi\). This setup had two main limitations: (i) bounded \(w\) does not imply bounded \(\xi\), so the link with classical robust control (as exploited for additive noise in [57]) is lost; (ii) the setup is inapplicable when only state measurements are available. This section resolves these issues.

A. MEASUREMENT MODEL

Considering the dynamics (13), the goal is to estimate the true \(W_c := E[ww^\top]\). This should be accompanied by a similar concentration inequality as in Lemma 5.3 without assuming direct access to samples \((w_i)_{i=1}^N\). Instead, we measure a sequence of \(N\) independent samples \((x_{i+1}, z_i)_{i=1}^N\), with \(z_i = (x_i, u_i)\) the augmented state, satisfying the measurement model

\[
x_{i+1} = [\mathcal{M}; I_{n_x}, z_i, u_i]
\]

where \((w_i)_{i=1}^N\) denotes the noise sequence.

There are two important properties that (20) should satisfy: (i) the “span” of the (mixed) model tensor \(\mathcal{M}\) should describe the true dynamics (i.e., Assumption 1); and (ii) the data should be sufficiently rich to render \(W_c\) observable. This final property is akin to persistency of excitation. We specifically assume that the support of \(z_i\) is not degenerate, as is formalized in the following assumption.

Assumption 2 (Nondegenerate support): We assume that \((z_i)_{i=1}^N\) is a sequence of i.i.d. copies of a random vector \(z\), with a distribution that is dominated by the Lebesgue measure.

A direct consequence of Assumption 2 is that, if \(N \geq n_z\)

\[
P[\text{rk}(z_1, \ldots, z_N)] = n_z
\]

(21)

Remark 5.1: Independence is guaranteed for each \(i\) by either (i) taking a known random \(z_i = (x_0, u_0)\), updating (1) once and taking \(x_{i+1} = x_1\); or (ii) taking a known \(x_0\), updating (1) \(T\) times for random \(u_t\) and taking \(x_{i+1} = x_T\) and \(z_i = (x_{T-1}, u_{T-1})\).

We refer to (i) as random initialization and to (ii) as rollout similar to [8] and [42].

For rollout, it is difficult to verify (21). Usually, the notion of controllability and persistency of excitation is employed. Full generalizations of these concepts for multiplicative noise do not exist to our knowledge. Note, however, that the event in (21) is related to the zeros of a polynomial. So an argument similar to the one in Proposition 5.2 below is applicable. In

\[^{3}\text{We will henceforth denote the true values using a star subscript.}\]

\[^{4}\text{More details on Assumption 2 are given in [56].}\]
practice one can sample \( u_t \) from a distribution dominated by the Lebesgue measure and test whether (21) holds for \( N = n_z \). When taking more samples, the validity of (21) will stay the same with probability one.

### B. Noise Observability

In some cases, the disturbances can be inferred exactly from state measurements. Specifically, in (20), we can expand the definition of the Tucker product as in Remark 3.10 and plug in the definition of \( \tilde{x}_2 \) to get

\[
x_{i+1} = (\mathcal{M} \otimes_z z_i) \otimes_x w_i = \left( \sum_{j=1}^{n_x} z_{ij} M_{j:j} \right) w_i.
\]  

(22)

Thus, if \((\mathcal{M} \otimes_z z_i)\) is left invertible, we can uniquely identify \((w_i)_{i=1}^N\) from \((x_{i+1}, z_i)_{i=1}^N\). Invertibility of \((\mathcal{M} \otimes_z z_i)\), however, is analogous to invertibility of a linear subspace of matrices (or linear forms), which is an unsolved problem in general (cf., [68] for 3 × 3 matrices). We have the following.

**Proposition 5.2:** Let \( z \) denote a random vector with non-degenerate support (cf., Assumption 2) and \( n_x \geq n_w \). Then, \((\mathcal{M} \otimes_z z)\) is left invertible either with probability one or zero. **Proof:** A high-level proof is given in [69, p. 2] in the setting of invertibility of linear forms.

This suggests that it is often sufficient to sample \((\mathcal{M} \otimes_z z)\) for a random \( z \) and check its invertibility. The conclusion will then generalize to other realizations, except for a set of measure zero. When invertibility holds, (22) can be solved for \( w_i \) exactly. Then the following holds.

**Lemma 5.3:** Let \( \mathcal{W} = \{ w \in \mathbb{R}^{n_w} : \|w\|_2 \leq r_w \} \). Assume we have a set of i.i.d. samples \((w_i)_{i=1}^N\) of a random vector \( w \) and let \( \hat{W} := \sum_{i=1}^N w_i w_i^\top/N \). Then

\[ \text{IP}\left[ \|\hat{W} - W\|_2 \leq \beta W \right] \geq 1 - \delta \]

when \( \beta W = \frac{r_w^2}{\sqrt{2N}\ln(2n_w/\delta)}/N \).

**Proof:** The result follows from a direct application of Lemma 2.1 by noting that the terms in the error satisfy \(-r_w^2 I \leq -w_i w_i^\top \leq W - w_i w_i^\top \leq W' \leq r_w^2 I \).

We also want to estimate \( W \) in the other case (i.e., when \( n_z < n_w \)). For example, for a model-free basis, we have \( n_w = n_x n_z > n_x \). To do so, we design least squares estimators.

### C. Least Squares Estimator

To design a least square (LS) estimator, we need to convert (20) into an expression linear in \( w_i w_i^\top \). We can do so by using Lemma 3.15 on (20) and use \( I_{n_x} \otimes \hat{I}_{n_x} = I_{n_x} \), which gives

\[
x_{i+1} \otimes_x x_{i+1} = \left[ \mathcal{M} \otimes_z \mathcal{M}; I_{n_x} \otimes z_i \otimes_x z_i, w_i \otimes_x w_i \right].
\]  

(23)

Note \( E[w_i \otimes_x w_i] = E[\text{vec}(w_i w_i^\top)] =: \text{vec}(W_*) \). We can write this in terms of linear equations by expanding the Tucker operator (cf., Definition 3.9 and Remark 3.10)

\[
x_{i+1} \otimes_x x_{i+1} = (\mathcal{W} \otimes_z (z_i \otimes z_i)) \text{vec}(W_*)
\]

\[
+ (\mathcal{W} \otimes_z (z_i \otimes z_i)) \eta_i
\]

with \( \mathcal{W} = \mathcal{M} \otimes_z \mathcal{M} \) and noise \( \eta_i = w_i \otimes x_i - \text{vec}(W_*) \). As usual least squares, stacking the equations gives

\[
Y_N = Z_N \text{vec}(W_*) + E_N
\]  

(24)

where we stack that data as follows:

\[
Y_N = (x_1 \otimes x_1, \ldots, x_{N+1} \otimes x_{N+1})
\]

\[
Z_N = \left[ \mathcal{W} \otimes_z (z_1 \otimes z_1); \ldots; \mathcal{W} \otimes_z (z_N \otimes z_N) \right]
\]

(25a)

\[
E_N = \left[ [\mathcal{W} ; z_1 \otimes z_1; \eta_1]; \ldots; [\mathcal{W} ; z_N \otimes z_N; \eta_N] \right].
\]

(25b)

Note that \( E[\eta_i] = 0 \) so \( E[E_N] = 0 \), which suggests the LS estimate

\[
\text{vec}(\hat{W}) = Z_N^\top Y_N.
\]  

(26)

A similar estimator to \( \hat{W} \) is used in compressed covariance sensing [51], where it is observed that it acts as a good heuristic for the maximum likelihood estimator when \( w \) is Gaussian. Noting that each element of the vector equation (23) constitutes a bilinear form of \( z_i \), reveals a connection with bilinear estimation [52]. There, adjusted LS estimators exist that are consistent even when \( z_i \) is perturbed by random noise. For simplicity, we consider exact state measurements here, which enables the use of ordinary LS, for which it is convenient to derive concentration inequalities.

### D. Error Analysis

We analyze the error of the LS estimator (26) under Assumptions 1–2. The discussion differs somewhat from the classical case due to the biased estimate (26). However, this bias is inconsequential as our focus is on estimating the second moment dynamics. We demonstrate that the estimate \( \mathcal{E}(\hat{W}; \cdot) \) captures the second moment dynamics \( \mathcal{E}_w := \mathcal{E}(W_\cdot; \cdot) \) without bias. This phenomenon is commonly observed in the system identification of multiplicative noise, as noted in [43] and [45]. Since this complicates the error analysis, we provide only the intuition here, with formal statements deferred to Appendix C.

The overall estimation error is given as

\[
\text{vec}(\hat{W} - W_*) = (I - Z_N^\top Z_N) \text{vec}(W_*) + Z_N^\top E_N.
\]  

(27)

We first discuss the nuances associated with the first term on a high level, before formally stating the error bound.

**Estimator bias:** Noting that the second term is zero mean, the first term in (27) constitutes the bias in the estimate. This bias is characterized exactly in Lemma 3.1 and is often nonzero. The reason for this can be given in terms of the matrix associated with \( \mathcal{E}(W; \cdot) \) denoted as \( \mathcal{E} \in \mathbb{R}^{n_w \times 3n_x} \), which uniquely determines the second moment dynamics. Our procedure parametrized \( E \) in terms of \( W \in \mathbb{R}^{n_w} \). In other words, we use \( d_{n_x} \) parameters to describe \( d_{n_x} \) degrees of freedom. In the model-free case (i.e., when \( n_w = n_x n_z \)) we would always overparametrize \( \mathcal{E} \) for \( n_x > 1 \).

One might therefore argue that \( E \) should be parametrized directly. We instead use \( W \) for three reasons: (i) often a description like (1) with a bound on \( \|w\|_2 \) is more natural in practical applications and such bounds do not translate well into bounds on \( E \); (ii) control synthesis using a confidence set over \( W \) is
more tractable; and (iii) the bias is inconsequential for the LQR cost and stability analysis.

Here, (iii) is caused by the fact that the bias always lies in the kernel of the second moment dynamics [cf., (15)] \( E(\tilde{W}; \tilde{Z}) \) w.r.t. \( W \). This kernel, in fact, is also the cause of the bias in the first place as formalized in Lemma 3.2.

**Data-driven error:** We are now ready to state a data-driven bound, suitable for control synthesis. To do so, we first consider some auxiliary operators. Letting \( \mathcal{M} \) denote the model tensor (cf., Assumption 1) and \( \mathcal{W} := \mathcal{M} \otimes_a \mathcal{M} \) as before, we introduce

\[
\mathcal{W}(z z^\top) := (\mathcal{W} \times_2 (z \otimes_a z))(\mathcal{W} \times_2 (z \otimes_a z))
= \mathcal{W}(3)((z z^\top \otimes a z z^\top) \otimes I_{n_a}) \mathcal{W}^\top(3)
\]

where the equality is shown in Lemma (i). Moreover, let

\[
\mathbf{H}_i := \left[ \sum_{j=1}^{N} \mathcal{W}(z_j z_j^\top) \right]^\top \mathcal{W}(z_i z_i^\top)
\tag{28}
\]

with \( \mathbf{H}_i \in \mathcal{S}_N \), the operator associated with \( \mathbf{H}_i \) has norm \( \| \mathbf{H}_i \| \leq \sqrt{\tau_W n_W} \| \mathbf{H}_i \|_2 \) (cf., Lemma 3.3). One can show that (cf., Corollary 3.4) the second term in (27) equals

\[
\text{vec}_a \left( \sum_{i=1}^{N} \mathbf{H}_i(w_i w_i^\top - W_a) \right)
\]

which is an i.i.d. sum of bounded random matrices. Hence, a matrix Hoeffding bound is applicable:

**Theorem 5.4:** Let \( \mathbb{E}_s \) denote the true moment dynamics as in (15). Assume access to samples \( (x_{i+1}, z_i) \) satisfying (20) and Assumption 2 and a model tensor \( \mathcal{M} \) satisfying Assumption 1.

If \( \tilde{W} \) is the LS estimate (26) and \( N \geq a_n \), then

\[
\mathbb{P}[\mathbb{E}(\tilde{W}; Z) \leq \mathbb{E}_s(Z) \leq \mathbb{E}(\tilde{W}; Z)] \geq 1 - \delta
\]

for all \( Z \in \mathbb{S}^n \), with \( \tilde{W} := \tilde{W} - \beta_W I_{n_W} \) and \( \tilde{W} := \tilde{W} + \beta_W I_{n_W} \) where \( \beta_W = r_W^2 \sqrt{2 \log(2 n_W/\delta)} \) and \( \zeta_W^2 := \sum_{i=1}^{N} \| \mathbf{H}_i \|^2_2 \), where \( \mathbf{H}_i \) denotes the operators associated with \( \mathbf{H}_i \) in (28).

**Proof:** Deferred to Appendix C.

**Sample complexity:** Due to independence, one may expect the matrices \( \mathcal{W}(z_i z_i^\top) \) to behave similarly. Hence, one can expect \( \| \mathbf{H}_i \|^2_2 \) to decrease with \( 1/N \). So \( \zeta_W^2 = \sum_{i=1}^{N} \| \mathbf{H}_i \|^2_2 \) would decrease with \( 1/N \) too.

**Lemma 5.5:** Following the setting of Theorem 5.4 and assuming additionally that \( \| z_i \|_2 \leq r_z \) a.s. for all \( i \in \mathbb{N}_{1:N} \). Then, if \( N \geq \| \mathcal{W} \|^2_\tau_W \sqrt{r_W}/\tau_W \)

\[
\mathbb{P}\left[ \zeta_W \leq \frac{\sqrt{n_W} \| \mathcal{W} \|^2_2 \sqrt{2 \ln(2 n_W/\delta)}}{\sqrt{N} \gamma_W - \| \mathcal{W} \|^2_\tau_W} \right] \leq \delta
\]

with \( d_W = r_W(\mathcal{W}(3)) \), \( \tau_W := \sqrt{2 \ln(2 d_W/\delta)} \) and

\[
\gamma_W := \lambda_{d_W} \left( \mathcal{W}(3) \left( E[z z^\top \otimes a z z^\top] / r_z^4 \otimes I_{n_a} \right) \mathcal{W}(3) \right)
\]

VI. DISTRIBUTIONALLY ROBUST CONTROL SYNTHESIS

In this section, we design approximate reformulations of (LQR). The goal is to preserve the properties of Theorem 4.7. Specifically, we want to synthesize a distributionally robustly stabilizing controller, by solving a SDP as in Theorem 4.9. The synthesis procedure is described in Section VI-A. We also evaluate the sample complexity in Section VI-B.

**A. Synthesis**

In this section, we use our confidence bound in Theorem 5.4 to synthesize a controller that stabilizes the true system with high probability. Similarly, to Section IV-C, we do so by investigating (LQR). Specifically, consider the CP version of (LQR)

\[
\min_{\tilde{W} \succeq W} \max_{\tilde{W} \succeq W} \sum_{t=0}^{\infty} \text{Tr}(Z_t H_t).
\]

subject to \( Z_t = [X_t, V_t, V_t^\top, U_t] \) subject to (54). The constraints on \( W \) define the equivalent of the ambiguity set in (LQR). Next, we solve (LQR) by generalizing Theorem 4.7. We will use the closed-loop policy \( \Pi_K \) as defined in (18) and

\[
\mathbb{E}_K(W; \cdot) := \mathbb{E}(W; \Pi_K(\cdot)).
\]

We have the following.

**Theorem 6.1:** Assume \( H = \text{blkdiag}(Q, R) \succ 0 \) and

\[
\exists K : \rho(\mathcal{E}(K(W; \cdot))) < 1, \forall W : W \preceq W \preceq \tilde{W}.
\]

Then, the policy \( Z_t = \Pi_K(X_t) \) is optimal for (LQR), with \( (\bar{P}, \bar{K}) \) the optimal solution as described in Theorem 4.7, where we assume \( \mathcal{E} = \mathcal{E}(W; \cdot) \). Moreover

1. \( \rho(\mathcal{E}(K(W; \cdot))) < 1, \forall W : W \preceq W \preceq \tilde{W} \); and
2. \( \mathcal{E}(\mathcal{LQR}_{\text{cp}}) = \text{Tr}[\bar{P} \cdot X_0] \).

**Proof:** The proof is similar to that of [14, Prop. 9] for multiplicative noise, and is based on showing that the Bellman operator is of the same structure as in Theorem 4.7 by using Corollary 3.14. A full proof for the CP case is given in [56].

Stabilizability is replaced by a robust equivalent. When violated, (LQR) is infeasible and the SDP in Theorem 4.9 is unbounded. If this occurs, the user should gather more data or verify if the nominal system is stabilizable using first principles (e.g., whether the uncertainty threshold is exceeded [34]). When \( \mathcal{E}(W; \cdot) \) is not stabilizable [i.e., (19) is unbounded], then the true system will not be either with high probability. In that case, not much can be done from a control perspective besides redesigning actuators.
Theorem 5.4 implies that the true \( \mathcal{E} \), satisfies \( \mathcal{E}(W) \leq \mathcal{E}(\bar{W}) \) with probability at least 1 − δ. So, we have:

**Corollary 6.2:** Let \( \bar{K} \) denote the DR policy as in Theorem 6.1 and assume the setup of Theorem 5.4 holds. Moreover, let \( \bar{\mathcal{E}} \) denote the true moment dynamics.

Then with probability at least 1 − δ

i) \( \rho(\mathcal{E}(\Pi_K(\cdot))) < 1 \);

ii) \( \mathcal{E}(\bar{\mathcal{L}}_{\mathcal{R}}) \geq \mathcal{E}(\mathcal{L}_{\mathcal{R}}) = \mathcal{V}_{\mathcal{L}}(\mathcal{L}_{\mathcal{R}}) \).

**Proof:** By Theorem 5.4, \( \mathcal{E}(W) \leq \mathcal{E}(\bar{W}) \) with probability at least 1 − δ. Thus, \( \exists \bar{W} \); \( W \leq \bar{W} \leq \bar{W} \) such that \( \mathcal{E}(W) = \mathcal{E}(\bar{W}) \). So stability (i) follows from Theorem (i) and (ii) follows by maximizing the inequality over \( \bar{W} \) for the inequality and Theorem 4.5 for the equality.

**Remark 6.3:** As in Theorem 4.5, we have \( \mathcal{V}_{\mathcal{L}}(\bar{\mathcal{L}}_{\mathcal{R}}) = \mathcal{V}_{\mathcal{L}}(\mathcal{L}_{\mathcal{R}}) \). To save space, we omit a proof and instead claim only \( \mathcal{V}_{\mathcal{L}}(\bar{\mathcal{L}}_{\mathcal{R}}) \geq \mathcal{V}_{\mathcal{L}}(\mathcal{L}_{\mathcal{R}}) \). This is sufficient for safety critical applications where an upper bound is sufficient for stability. We show consistency of \( \bar{\mathcal{L}}_{\mathcal{R}} \) later to further strengthen the result.

**B. Sample Complexity**

As shown in the previous section, the solution of the DR problem in Section V is described by a perturbed Riccati equation. As such, we use perturbation analysis to bound the suboptimality of the DR controller when applied to the true dynamics. We leverage a previous result by the authors in [55] to do so.

The final complexity bound is as follows:

**Theorem 6.4:** Assume Assumptions 1 and 2 (for \( \|z_i\|_2 \) bounded a.s.) and let \( \bar{K} \) be the optimal controller in Theorem 6.1. Then, for sufficient samples \( N \) and with probability of at least 1 − δ

\[
\operatorname{Tr} \left[ \sum_{t=0}^{\infty} Z_t H \right] - \mathcal{V}_{\mathcal{L}}(\mathcal{L}_{\mathcal{R}}) = O\left( \frac{1}{N} \right)
\]

where the first term is the cost of \( \mathcal{L}_{\mathcal{R}} \) achieved for \( Z_t = \Pi_{\bar{K}}(X_t) \) with true moment dynamics \( X_{t+1} = \mathcal{E}(Z_t) \).

Moreover, by Theorem 4.5, for sufficient samples \( N \) and with a probability of at least 1 − δ

\[
\mathbb{E} \left[ \sum_{t=0}^{\infty} x_t^\top Q x_t + u_t R u_t \right] - \mathcal{V}_{\mathcal{L}}(\mathcal{L}_{\mathcal{R}}) = O\left( \frac{1}{N} \right)
\]

where \( x_{t+1} = A(x_t)x_t + B(v_t)u_t \) as in (1), \( u_t = \bar{K}x_t \). So, the first term is the cost achieved when applying \( \bar{K} \) to the true multiplicative noise dynamics (1).

**Proof:** We suggest some modifications of the proof of [55, Cor. III.4] to show the second result. The first result then follows directly from Theorem 4.5.

**Modifications:** Assumption II.2] assumes a relative bound like \( \alpha W_t \leq W_t \leq \alpha W_{\bar{t}} \), which is \( \alpha \leq \bar{\alpha} = \alpha(1/\sqrt{N}) \) with high probability, \( W_t \), \( \alpha \) and \( \bar{\alpha} \) some estimate. Instead, we have \( \leq \bar{\alpha} W_t \leq 2\alpha W_t \) with \( \bar{\alpha} = O(1/\sqrt{N}) \) by Theorem 5.4 and Lemma 5.5 (where we ignore the effect of the bias, without loss of generality due to Lemma 3.2). So, the proofs need to be modified. Specifically, we replace [55, Lemma IV.3] with a bound like

\[
\left\| A_{(2)}(\bar{W} - W_t \otimes P)A_{(2)}^\top \right\|_2 \leq 2\bar{\alpha} W_t \left\| A_{(2)}(I \otimes P)A_{(2)}^\top \right\|_2.
\]

The proof is analogous and is based on Corollary 3.14. Following the proof of [55, Lemma IV.3] and its dependencies, clearly only the constants in [55, Table 1] are affected. The rate shown in [55, Cor. III.4] remains the same. The full proof with explicit constants is given in [56, Appendix F].

**Remark 6.5:** Since the estimation error bounds are of the same order, we get the same sample complexity in the certainty equivalence setting [11, §3.1], where we use \( \bar{W} \) instead of \( W \). The proof is analogous.

In both cases the rate in terms of \( N \) is the same as in the additive noise case [9], [71]. This implies that, qualitatively in terms of samples, learning additive noise or multiplicative noise is similar. The dimensional dependency of the error, however, is \( n_2^2(n_x + n_u) \) for \( n_x \geq n_u \) and omitting logarithmic terms. This is worse compared to additive noise [9], [71], where a lower bound is established of \( \sqrt{n_2^2 n_u} \). Note that our bound is not a lower bound, so further research might lead to a tighter dimensional dependency. Exploiting \( \bar{W} \) in Theorem 5.4 could be a first step.

**VII. NUMERICAL EXPERIMENTS**

In this section, we numerically investigate the methods developed in this article. We begin with the simple case of repeated initialization (cf., Remark 5.1), comparing both the case where the modes are known and where they are not, i.e., the model-free setting. We show how, in the first setting, our bounds are sufficiently tight to enable control synthesis. Next, we use data generated using rollout to produce similar estimates and we highlight the differences and challenges associated with doing so. We also show that experimentally, our method also works when using only a single trajectory, although theoretical guarantees are not yet available in this setting. In the technical report [56], we compare with the averaging rollouts approach in [42] and [43], and show how structural information can be exploited to get tighter estimates of the uncertainty.

**A. Accuracy of the Estimate and Bounds**

We consider the case \( n_x = 2 \) and \( n_u = 1 \) with modes

\[
A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
\]

and \( B_1 = B_2 = 0 \). The disturbance is sampled uniformly from the ball \( \{w \in \mathbb{R}^3 : \|w - \mu\|_2 \leq 0.25\} \) with \( \mu = (0, 0, 0, 1) \).

**Repeated initialization:** To generate measurements, we sample a \( \omega_0 = (x_0, u_0) \) uniformly from an Euclidean ball of radius one and then propagate the dynamics by one step to get a measurement \( x_1 \), i.e., using repeated initialization as in Remark 5.1. From this data, an estimate \( \bar{W} \) and associated radius \( \beta_W \) are determined as prescribed in Theorem 5.4.

To empirically quantify the accuracy of the scheme, the procedure above is repeated \( M = 100 \) times for each sample count \( N \). For each estimate, we compute \( \|\bar{W} - W_t\|_2 \) and \( \beta_W \). A confidence plot is provided in Fig. 2. The middle plot depicts the result when \( \mathcal{M} \) is selected, based on true mode info (i.e., \( \mathcal{M}(i) = I \)).
Empirical result for toy problem using rollout data: (left) when using true modes in $\mathcal{M}$; (right) when using no mode information, i.e., model-free. The dashed line depicts the radius predicted by Theorem 5.4. Each colored area is a 0.1 confidence interval surrounding the median.

We determined $W_*$ in the first setting by noting that $E[w w^\top] = r_w^2 I_{n_w} / (n_w + 2)$ when $w$ is sampled uniformly from $\{ w \in \mathbb{R}^{n_w} : \|w\|_2 \leq r_w \}$ — which can be verified by symmetry and solving a simple integral—and applying the appropriate transformation. Similarly, $W_*$ for the model-free case is then recovered by applying the transformation in (14).

The left-most plot depicts the error when $W$ is estimated directly from measurements of $w$ and the bound is as is in Lemma 5.3. It is clear from the figure that the empirical error does not increase much when we do not directly observe the disturbance—at least for this simple model. Instead, the main loss is in the bound, which is about an order of magnitude looser than the direct sample case. This is to be expected, however, and, as we confirm later using DR synthesis, the bounds are still practical in low-dimensional settings.

The middle and right-most plot depict the use of prior information and the model-free case, respectively. The fact that $W_*$ is not the same for both cases implies that the absolute errors are dissimilar. For a fair comparison, we additionally used the estimated second moment dynamics $\bar{E}$ as in (15). Specifically we plot $\| \bar{E} - E_* \|_2 / \| E_* \|_2$ where $\bar{E}$ is the matrix associated with the estimated dynamics and $E_*$ its true value. These are computed using Corollary 3.16. Note that $E_*^\sigma$ is the same independent of the selected $\mathcal{M}$. Fig. 3 shows the result.

It is clear that, for the low-dimensional example considered here, exploiting prior mode information has no significant advantage when it comes to estimation accuracy. Instead, the main advantage is a smaller ambiguity set. When using the model-free approach, the control synthesis problem remains infeasible for any tested sample count. That is, the ambiguity set exceeds the minimum size that can be stabilized by a single controller. When exploiting the prior information encapsulated in the modes, this is not the case.

Control synthesis: We can quantify the size of the ambiguity set through controller synthesis as in Theorem 6.1. The performance of a gain $\hat{K}$ is measured through the infinite horizon cost achieved on the true system. This cost equals the optimum of the following SDP, by Proposition 3.8 and Lemma 1.1

$$\min \{ \text{Tr}[P X_0] : P - \Pi^\text{K}_K (E_\sigma^*(P)) \succeq \text{blkdiag}(Q, R), \ P \succeq 0 \}.$$ 

We select $X_0 = I$ here and in Theorem 4.9 for control synthesis. Also, let $Q = I$ and $R = 10$. The relative error with the true optimum of ($LQR_\sigma$) is then a metric for the accuracy of the ambiguity set. The result is depicted in Fig. 4. Note that the rate is as predicted in Theorem 6.4.

The dashed line in Fig. 4 depicts the cost for the trivial ambiguity, which only uses $\|w\|_2 \leq r_w$ implying $\|W_*^\sigma\|_2 \leq r_w^2$. This constraint is also used to synthesize a controller with Theorem 6.1. We say that the learned ambiguity set is informative when the performance of the associated controller improves upon the trivial one.

Rollout: Repeated initialization is not realistic in practice, since it assumes we can directly select all states of the system. Rollout instead initializes the system at some—easily realizable—initial state and then applies a sequence of control actions to excite the system. We use the state $x_0 = (1, 1)$ and use the control law $u_t = [-0.5, -0.2] x_t + \delta_t$ with $\delta_t$ sampled uniformly from a Euclidian ball of radius 35 at each time step to integrate the dynamics for $T = 25$ time steps. Then $z_i = (x_{T-1}, u_{T-1})$ and $x_{i+1} = x_T$ are used as data points (cf., Remark 5.1) for $i = 1, \ldots, N$. Here, $N$ is the amount of rollouts, which we also refer to as the sample count.

Similarly to before, we can evaluate the accuracy of the estimate empirically, by resampling data sets $M = 100$ times and computing the errors. The result is depicted in Fig. 5. Comparing with Fig. 2, we see, as expected, that estimation is more challenging when using rollout data. This is likely explained

---

6To shift the ball, we shift $\bar{w} = w + \mu$. Hence, $E[\bar{w} \bar{w}^\top] = W_\mu + \mu \mu^\top$. 

---

Fig. 2. Error in moment estimation using repeated initialization: (left) when directly observing the disturbance; (mid) when using the true modes in $\mathcal{M}$; (right) when using no mode information, i.e., model-free. The dashed line depicts the radius predicted by Theorem 5.4. Each colored area is a 0.1 confidence interval surrounding the median.

Fig. 3. Estimation accuracy of matrix of second-order dynamics and comparison between model-free and true mode based estimation. Each colored area is a 0.1 confidence interval surrounding the median.

Fig. 4. Evaluation of suboptimality of DR control synthesis. The colored area is a 0.1 confidence interval surrounding the median.

Fig. 5. Empirical result for toy problem using rollout data: (left) using the tail of $\#n$ samples rollouts of length 25, with predicted radius as a dashed line; (right) using one rollout of length $\#n$ samples. Each colored area is a 0.1 confidence interval surrounding the median.
by the distribution of \( z \) being less suitable for identification as discussed in Remark 5.7.

**Single trajectory:** The right plot in Fig. 5 also depicts the empirical error when we simply use one rollout of length \( T = N \). So the data is then \( (x_i, u_i) \) and \( x_{i+1} \) for \( i \in \mathbb{N}_{0:N-1} \). It is clear that the estimate becomes more accurate when the trajectory length is increased, showing—at least empirically—that our approach can also work for single trajectory identification. This was also observed for a similar setup in [45], yet a convergence proof and associated conditions on the exciting inputs are still unavailable (cf., Remark 5.1).

### VIII. Conclusion

We developed a novel system identification scheme for linear dynamics with state- and input-multiplicative noise. The resulting estimators are shown to converge at a \( 1/\sqrt{N} \) rate, with \( N \) the number of samples, when the data is generated either using rollout or through repeated initialization of the dynamics. We illustrated empirically that—for simple dynamics—the constants in our bounds are practical, i.e., they can be used for DR control synthesis. Moreover, the tightness of the bound is similar to that of a usual matrix Hoeffding bound. Also, the DR control synthesis problem was shown to converge at a \( 1/N \) rate to the true optimum.

For more complex dynamics, it is likely that the bounds are not sufficiently tight to enable DR control synthesis. In that case, either bootstrapping can be exploited as in [8] and [72] or the DR scheme can be used to identify the robustness of the certainty equivalent controller as in [32].

From these tests, it is clear that the correlation between errors does not accumulate when using single trajectory data. A formal proof of this fact and a sample complexity analysis, however, proved challenging and is considered an open problem for future work.

### Appendix A

**Linear Quadratic Regulation**

This section gives proofs related to LQR. We begin by introducing some auxiliary operators and their adjoints.

**Lemma A.1:** Let \( \mathcal{E} \) be as in (15), \( \Pi_K(X) = [I; K]X[I, K^\top] \), \( \mathcal{E}_K(X) = \mathcal{E} (\Pi_K(X)) \). Then, we have the following adjoints:

i) \( \Pi^*_K(H) = [I, K^\top]H[I, K] \);

ii) \( \mathcal{E}^*_K(P) = \begin{bmatrix} \mathcal{F}^*(P) & \mathcal{H}^*(P) \\ \mathcal{H}^*(P) & G^*(P) \end{bmatrix} \);

iii) \( \Pi^*_K(E^*_K(P)) = \mathcal{E}_K^*(\mathcal{E}^*_K(P)) \).

The elementary adjoints are \( \mathcal{F}^*(P) = \mathcal{A}_2(W \otimes P)\mathcal{A}_2^\top \), \( \mathcal{H}^*(P) = \mathcal{B}_2(W \otimes P)\mathcal{B}_2^\top \) and \( G^*(P) = \mathcal{B}_2(W \otimes P)\mathcal{A}_2^\top \).

**Proof:** Only elementary algebra, Lemma 3.12 and Corollary 3.16 are used.

**Lemma A.2:** Any \( Z \succeq 0 \) can be partitioned as

\[
\begin{bmatrix}
X & X K^\top \\
K X & K X K^\top + \Delta
\end{bmatrix}
\]

for \( \Delta \succeq 0 \) and \( X \succeq 0 \).

**Proof:** Let \( Z = [X, V; V^\top, U] \). By [73, §A.5.5] \( Z \succeq 0 \) iff

\[
X \succeq 0, \quad imV \subseteq \text{im}X, \quad U - V^\top X^\top V \succeq 0.
\]

So there exists a \( K \) s.t. \( V = X K^\top \) and \( \Delta = U - V^\top X^\top V = U - K^\top X K \succeq 0 \).

**Proof of Theorem 4.5:** We use Assumption 1 and Lemma 4.1 to consider (13) instead of (1) without loss of generality. We start with a feasible sequence for (LQR) and construct an (equivalent) feasible sequence for (LQRP). Let \( H = \text{blkdiag}(Q, R) \), \( Z_t = E[z_tz_t^\top] \geq 0 \) with \( z_t = (x_t, u_t) \) and \( x_t, u_t \) satisfying (13).

Note that, by (15), any sequence \( Z_t \) constructed as such satisfies the constraints of (LQRP). Then

\[
E \left[ \sum_{t=0}^{\infty} x_t^\top Q x_t + u_t^\top R u_t \right] = \sum_{t=0}^{\infty} \text{Tr}[H Z_t].
\]  

So the cost of \( \{x_t, u_t\} \) for (LQRP) equals that of the constructed \( Z_t \) in (LQRP). So \( \text{val}(LQRP) \leq \text{val}(LQR) \).

Next, we argue that for any sequence \( Z_t \) feasible for (LQR), we can construct a sequence \( z_t = (x_t, u_t) \) feasible for (LQRP). Note that for any \( x_0 \geq 0 \) we can easily construct a random vector \( x \) with \( E[x x^\top] = X \). So let \( x_0 \) be such a vector for \( X_0 \). Then, by Lemma 1.2, \( Z_0 \) can be partitioned as \( [X_0, X_0 K_0; K_0 X_0, K_0 X_0 K_0] + \Delta_0 \) with \( \Delta_0 \succeq 0 \) for which we pick a random vector \( \delta_0 \). So if we take \( u_0 = K_0 x_0 + \delta_0 \) and \( z_0 = (x_0, u_0) \) then \( E[z_0 z_0^\top] = Z_0 \). Repeating the same argument starting at \( x_1 \) and continuing for all time steps allows us to construct a feasible trajectory for (LQRP). Again by \( E[z_t z_t^\top] = Z_t \) and (29) the cost for both trajectories is the same. Thus, \( \text{val}(LQRP) \geq \text{val}(LQR) \).

### Appendix B

**Vector and Matrix Concentration**

**Lemma B.1:** For some \( \gamma \in \mathbb{R}^N \) and a random, independent sequence \( \{X_i\}_{i \in \mathbb{N}_0} \subset \mathbb{S}_d \) with \( \|X_i\| \leq \gamma \) a.s. and \( E[X_i] = 0 \). Then, for \( \beta \geq 0, Y = \sum_{i=0}^N X_i \)

\[
\mathbb{P}[\gamma_{\max}(Y) \geq \beta] \leq \delta; \quad \mathbb{P}[\gamma_{\min}(Y) \leq -\beta] \leq \delta
\]

for \( \delta = d \exp(-\beta^2/2\gamma^2) \). Moreover, \( \mathbb{P}[\|Y\|_2 \geq \beta] \leq 2\delta \).

**Proof:** We extend the proof of [57, Lemma II.2]. Let \( X_i = V_i^L \Lambda_i V_i^L \) be the eigenvalue decomposition. Hence

\[
E[\exp(\theta X_i)] = E[V_i^L \exp(\theta \Lambda_i) V_i^L^\top]
\]

\[
\leq E \left[ \gamma_i I + X_i e^{\theta \gamma_i} + \frac{\gamma_i I - X_i}{2\gamma_i} e^{-\theta \gamma_i} \right]
\]

by convexity. Also from \( E[X_i] = 0 \) we have, \( E[\exp(\theta X_i)] \leq \cosh(\theta \gamma_i) I \leq \exp(\theta^2 \gamma_i^2/2) I \).

Applying [53, Th. 3.6.1] results in

\[
\mathbb{P}[\gamma_{\max}(Y) \geq \beta] \leq \inf_{\theta > 0} d \exp(-\|\gamma\|_2^2/2 - \theta \beta)
\]

\[
\mathbb{P}[\gamma_{\min}(Y) \leq -\beta] \leq \inf_{\theta > 0} d \exp(-\|\gamma\|_2^2/2 + \theta \beta)
\]

Minimizing over \( \theta \) gives the required result for \( \lambda_1 \) and \( \lambda_d \). Taking a union bound gives the spectral norm bound.

**Lemma B.2:** Let \( \gamma \in \mathbb{R}^N \) and consider a random, independent sequence \( \{x_i\}_{i \in \mathbb{N}_0} \subset \mathbb{R}^d \) with \( \|x_i\| \leq \gamma \) a.s. Then, for
\( \beta \geq 0, y = \sum_{i=1}^{N} (x_i - E[x_i]) \)

\[
\mathbb{P}[\|y\| \geq 2\|\gamma\|_2 + \beta] \leq \delta
\]

for \( \delta = \exp(-\beta^2/2\|\gamma\|_2^2) \).

**Proof:** We use the techniques of [74, 6]. Introducing a filtration, let \( \mathcal{F}_t \) the \( \sigma \)-algebra generated by \( \{x_i\}_{i \in \mathbb{N}_{1:N}} \) with \( \mathcal{F}_0 \) the trivial algebra. Let \( E_i[x] = E[x | \mathcal{F}_i], \Delta E_i[x] = E_i[x] - E_{i-1}[x] \) and \( d_i = \Delta E_i[\|y\|_2] \). Then, by telescoping over \( i \), we have \( \sum_{i=1}^{N} d_i = \|y\|_2 - E[\|y\|_2] \). Noting that \( y - x_1 \) does not depend on \( x_1 \) and letting \( \delta_i = \|y\|_2 - \|y - x_1\|_2 \) gives

\[
\Delta E_i[\delta_i] = d_i - E_i[\|y\|_2 - \|y - x_1\|_2] + E_{i-1}[\|y - x_1\|_2] = d_i.
\]

Then, by the triangle inequality, \( |\delta_i| = \|y\|_2 - \|y - x_1\|_2 \leq \|y - x_1\|_2 \leq \|y\|_2 \). Letting \( e_i := E_i[\delta_i] \) then gives \( |d_i - e_i| = |E_i\delta_i| \leq |\delta_i| \).

Hence, we have constructed a sequence of zero-mean, independent random variables \( \{d_i\}_{i \in \mathbb{N}_{1:N}} \) that take their value a.s. in \( [-\|\gamma\|_2, \|\gamma\|_2] \). Applying a classical Hoeffding bound [75, Th. 2.8] gives

\[
\mathbb{P}\left[ \sum_{i=1}^{N} d_i \geq \beta \right] \leq \exp\left(-\beta^2/2\sum_{i=1}^{N} \|\gamma\|_2^2\right) \tag{30}
\]

where, as shown before, \( \sum_{i=1}^{N} d_i = \|y\|_2 - E[\|y\|_2] \). We introduce a sequence of independent random vectors \( \{v_i\}_{i \in \mathbb{N}_{1:N}} \) which is identically distributed to \( \{x_i\}_{i \in \mathbb{N}_{1:N}} \). Let \( E_i[x] \) denote the expectation conditioned on \( \{x_i\}_{i \in \mathbb{N}_{1:N}} \). Then, by Jensen’s inequality, we can bound \( E[\|y\|_2] \) as

\[
E\left[ \sum_{i=1}^{N} x_i - E_i[x_i] \right] \leq E\left[ \sum_{i=1}^{N} x_i - v_i \right]_2 \leq E\left[ \sum_{i=1}^{N} (x_i - v_i) \right]_2.
\]

Using a classical symmetrization argument, introducing independent Rademacher random variables \( \{\epsilon_i\}_{i \in \mathbb{N}_{1:N}} \) gives

\[
E[\sum_{i=1}^{N} (x_i - v_i)]_2 \leq 2E[\sum_{i=1}^{N} x_i e_i]_2.
\]

Applying Jensen once more, followed by a triangle inequality gives:

\[
E[\|y\|_2] = 2E\left[ \sum_{i=1}^{N} x_i e_i \right]_2 \leq 2\sqrt{E\left[ \sum_{i=1}^{N} x_i e_i \right]^2} \leq 2\sqrt{2\sum_{i=1}^{N} \|x_i\|_2^2} \leq 2\|\gamma\|_2.
\]

Plugging into (30) concludes the proof.

**APPENDIX C**

**IDENTIFICATION MODEL**

We begin by showing some auxiliary results.

**Lemma C.1:** For \( Z_N \) as in (25a) and \( N \geq \delta_{n_x} \)

\[
Z_N Z_N = (\mathbf{W}^{(3)})^T \mathbf{W}^{(3)}
\]

with probability one over data \( \{z_i\}_{i=1}^{N} \) satisfying Assumption 2.

**Proof:** We can rewrite (25a) as

\[
Z_N = ([z_1 \otimes z_1, \ldots, z_N \otimes z_N]^T \otimes I_{n_{x_{n}}}) \mathbf{W}^{(3)}
\]

by [\(\mathcal{T} \times x; \mathcal{T} \times y\) = \(\mathcal{T} \times [x, y]^T\) [cf., (8)] and Proposition 3.11.

Let \( Z_N := [z_1 \otimes z_1, \ldots, z_N \otimes z_N]^T \). Then by \( N \geq \delta_{n_x} \) and Assumption 2, \(r_k(z_N) = \delta_{n_x} \) w.p. 1. Hence, \( Z_N \otimes I \) is also left-invertible [58, Th. 4.2.15] w.p. 1.

For two matrices \( U, V \) of conformable dimensions, with \( U^T U = I \), we mimic the proof of [76, Th. 1].

\[
V^T V = V^T U^T U V \equiv V^T U^T ((U V)^T)^T = V^T U^T ((U V)^T)^T = V^T U^T (U V)^T
\]

where both (a) and (b) use \( A^T A^T = A^T \) (cf., [76, eq. (2)]).

Applying this to \( Z_N \) then gives the required result.

**Lemma C.2:** Take \( \text{vec}_s(W) = (I - (\mathbf{W}^{(3)})^T \mathbf{W}^{(3)}) \text{vec}_s(W) \) for all \( W \in S^{n_{x_{n}}} \). Consider \( \mathcal{E} \) as in (15). Then

\[
\mathcal{E}(W + \tilde{W}; Z) = \mathcal{E}(W; Z)
\]

for all \( W \in S^{n_{x_{n}}} \) and \( Z \in S^{n_{z}} \).

**Proof:** From [56, Prop. A.1(iii)] and Proposition 3.11 we have

\[
\text{Tr}[P \mathcal{E}(\tilde{W}; Z)] = \left[ \mathcal{M} \otimes \mathcal{M}; \text{vec}_s(P), \text{vec}_s(Z), \text{vec}_s(\tilde{W}) \right] = (\text{vec}_s(P) \otimes \text{vec}_s(Z)) \mathbf{W}^{(3)} \text{vec}_s(\tilde{W}).
\]

for all \( P \in S^{n_{x_{n}}} \) and \( Z \in S^{n_{z}} \). By assumption, and \( AA^T A = A \mathbf{W}^{(3)} \text{vec}_s(\tilde{W}) = \mathbf{W}^{(3)} (I - (\mathbf{W}^{(3)})^T \mathbf{W}^{(3)}) \text{vec}_s(W) = 0 \).

Therefore, \( \text{Tr}[P \mathcal{E}(\tilde{W}; Z)] = 0 \) for all \( P \) and \( Z \), which is only possible if \( \mathcal{E}(\tilde{W}; Z) = 0 \) for all \( Z \in S^{n_{z}} \) (otherwise we can take \( P = \text{unvec}(\epsilon_i) \), with \( \epsilon_i \) the canonical basis vector and \( i \) the index of the nonzero element in vec(\( \mathcal{E}(\tilde{W}; Z)) \).

Next, we derive properties of \( W \) and \( H_i \) defined in (28).

**Lemma C.3:** Let \( \mathcal{W}(Z) := \mathbf{W}^{(3)} ((Z \otimes z) \otimes I_{n_{x_{n}}}) \mathbf{W}^{(3)} \).

Then

i) \( \mathcal{W}(z z^T) = (\mathbf{W} \times_2 (z \otimes z)) \mathbf{W} \times_2 (z \otimes z) \);
ii) \( \sum_{i=1}^{N} \mathcal{W}(z_i z_i^T) = Z_N^T Z_N \);
iii) \( \|\mathcal{W}\|_2 \leq \text{sup}_{\bar{z}}\{\|\mathbf{W}(z z^T)\|_2 : \|z\|_2 \leq 1\} \leq \|\mathbf{W}\|^2_2 \) with \( Z_N \) as in (25a).

**Proof:** The first result (i) follows by using Proposition 3.11 to write

\[
(\mathbf{W} \times_2 (z \otimes z))^T \mathbf{W} \times_2 (z \otimes z) = \mathbf{W}^{(3)} (z \otimes z) \otimes I(I((z \otimes z)^T \otimes I)\mathbf{W}^{(3)}.
\]

Using \((A \otimes X)(B \otimes Y) = AB \otimes XY \) [59, Th. E.1.3] and [59, Lemma E.1.2] to argue \((z \otimes z)(z \otimes z)^T = (zz^T \otimes z z^T)^T \).
proves (i). We can show (ii) by using (31) and applying similar tricks. Finally, (iii) is shown by using (i) to argue
\[ \|W\|_2 = \sup_{z \in \mathcal{B}} \{ \| (W \tilde{x}_2 (z \otimes z)) \|_2^2 \} = \sup_{x, w, z \in \mathcal{B}} \{ (x^\top (W \tilde{x}_2 (z \otimes z)) w)^2 \} \] (32)
where we used the variational representation of the spectral norm for the second equality and with \( x \in \mathbb{R}^{n_a}, w \in \mathbb{R}^{n_w} \) and \( \mathcal{B} \) the unit Euclidean ball of generic dimension. The squared quantity can be rewritten by noting
\[ x^\top (W \tilde{x}_2 (z \otimes z)) w = [W : x, z \otimes z, w]. \]
Noting that \( (z \otimes z)^\top (z \otimes z) = \|z\|_2^4 \) implies that we can relax (32) to take the supremum over \( z' \in \mathcal{B}_{n_a} \) instead of over \( z \otimes z \).
So
\[ \|W\|_2 \leq \left( \sup_{x, w, z' \in \mathcal{B}} \{ \| W : x, z', w \| \} \right)^2 \]
which, plugging in the definition of the tensor spectral norm [70], implies (iii).

A direct consequence of Lemma 3.3 is that Corollary C.4: Given \( Z_N, E_N \) as in (25a) and (25b), respectively, and \( W \) as in Lemma 3.3. Then
\[ Z_N^\dagger E_N = \left[ \sum_{i=1}^N W(z_i z_i^\top) \right]^\dagger \left[ \sum_{i=1}^N W(z_i z_i^\top) \eta_i \right] \]
with \( H_i \) the linear operator with matrix \( H_i \) as in (28).

**Proof:** Note that
\[ Z_N^\dagger E_N = (Z_N Z_N^\dagger) Z_N^\dagger E_N. \]
Then, the first equality follows by application of Lemma (ii) to replace \( (Z_N Z_N^\dagger) \) and Lemma (i) to show that \( Z_N^\dagger E_N = \sum_{i=1}^N W(z_i z_i^\top) \eta_i \). The second equality follows by definition of \( \mathcal{H}_i \) and \( \eta_i = \text{vec}_a(W_i w_i^\top - W) \).

We are now ready to prove the data-driven bound.

**Proof of Theorem 5.4:** By the classical LS error (27)
\[ \text{vec}_a(W - W) = (I - Z_N^\dagger Z_N) \text{vec}_a(W) + Z_N^\dagger E_N \]
where we need Assumption 1 to imply (20) holds using Lemma 4.1. Note that, by Corollary 3.4 with \( H_i \) as in (28).

\[ \text{vec}_a(Z_N^\dagger E_N) = \sum_{i=1}^N H_i(w_i w_i^\top - W). \]

Note that \( -r^2 I \preceq W \preceq w_i w_i^\top - W \preceq w_i w_i^\top \preceq r^2 I \). Hence, by definition of \( \| H_i \|_2 \), the spectral norm of \( X_i := H_i(w_i w_i^\top - W) \) is bounded as \( \| X_i \|_2 \leq \| H_i \|_2 r^2 \). The matrices \( X_i \) are, therefore, bounded and i.i.d. (by Assumption 2). Applying Lemma 2.1 and solving for \( \beta \) shows
\[ \mathbb{P}\left[ \sum_{i=1}^N \| H_i(w_i w_i^\top - W) \|_2 \leq \beta \right] \geq 1 - \delta. \] (33)
We have by Assumption 2 and Lemma 3.1 that
\[ (I - Z_N^\dagger Z_N) \text{vec}_a(W) = (I - (W \tilde{z} (3))) W \tilde{z} (3) \text{vec}_a(W). \]

Applying Lemma 3.2, thus, implies
\[ \mathcal{E}(W - W; Z) = \mathcal{E}(\text{vec}_a(Z_N^\dagger E_N); Z). \]
The bound on \( \| \text{vec}_a(Z_N^\dagger E_N) \|_2 \leq \| \sum_{i=1}^N H_i(w_i w_i^\top - W) \|_2 \) in (33), therefore, proves the result.

We can also derive the sample complexity for the simplified setting where \( \| z \|_2 \leq r_z \).

**Proof of Lemma 5.5:** To deal with noninvertibility issues, we introduce \( U \in \mathbb{R}^{n_w \times d_2} \), a unitary matrix whose columns span \( \text{im}(W(3)) \). For all \( i \in \mathbb{N}_1 N \), we consider the i.i.d. random matrices \( Y_i = U^\dagger W(z_i z_i^\top) U \) and \( Y = \sum_{i=1}^N Y_i \).

By Assumption 2 and Lemma 3.1: \( \text{im}(Z_N) = \text{im}(W(3)) \). So \( \lambda_{\text{min}}(Y) = \lambda_{\text{min}}(U^\dagger Z_N U) = \lambda_{\text{min}}(Z_N U) \). Moreover, \( Y \succ 0 \). Similarly, \( \mathbb{E}([W(z z^\top)](3)) = W(3) \mathbb{E}([z \otimes z](z \otimes z)^\top) \otimes I) W(3) \), with \( \mathbb{E}[z \otimes z](z \otimes z) \sim 0 \) by Assumption 2.

So \( \lambda_{\text{min}}(E[Y_i]) = \lambda_{\text{min}}(E[W(z z^\top)]) = \gamma_{W} r_z^2 \). Therefore, Lemma 2.1 can be applied to show
\[ \mathbb{P}(\lambda_{\text{min}}(Y - E[Y]) \geq -\beta) \geq 1 - \delta \]
with \( \beta = r_z^2 \| Y \|_2 \sqrt{\gamma_{W}} \tau_{W} \). Hence \( \lambda_{\text{min}}(Y) \geq \lambda_{\text{min}}(E[Y]) - \beta = N \gamma_{W} r_z^2 - \beta \).

So, since \( Y^\dagger Y \geq 1 = \lambda_1(Y^\dagger) = (\lambda_{\text{min}}(Y))^\dagger \)
\[ \| Y^\dagger Y \|_2 \leq \gamma_{W} / \sqrt{\gamma_{W} - N \gamma_{W} r_z^2} \]
Using Lemma 3.3, we have
\[ \zeta_{W}^2 = \sum_{i=1}^N \| H_i \|_2^2 \leq N \max_i \| H_i \|_2^2 \]
where
\[ \| H_i \|_2^2 \leq n_w \| Y^\dagger Y \|_2^2 \]
\[ \leq n_w \| Y^\dagger Y \|_2^2 \leq n_w \| (Y^\dagger Y) r_z^2 \|_2 \| W \|_2^2 r_z^2. \]

Therefore, \( \| H_i \|_2^2 \leq n_w \| (N \gamma_{W} - \sqrt{\gamma_{W}} \| W \|_2^2 r_z) \|_2^2 \).

Multiplying by \( N \) then gives us the claimed bound for \( \zeta_{W} \).

**REFERENCES**

[1] B. Recht, “A tour of reinforcement learning: The view from continuous control,” Annu. Rev. Control Robot. Autom. Syst., vol. 2, no. 1, pp. 253–279, 2019. [Online]. Available: https://www.annualreviews.org/doi/10.1146/annurev-control-053018-032825

[2] L. Hewing, K. P. Wabersich, M. Menner, and M. N. Zeilinger, “Learning-based model predictive control: Toward safe learning in control,” Annu. Rev. Control Robot. Autom. Syst., vol. 3, no. 1, pp. 269–296, 2020.

[3] L. Brunke et al., “Safe learning in robotics: From learning-based control to safe reinforcement learning,” Annu. Rev. Control Robot. Autom. Syst., vol. 5, pp. 411–444, 2022.

[4] D. P. Bertsekas, Dynamic Programming and Optimal Control–Vol. 1, 3rd ed. Nashua, NH, USA: Athena Scientific, 2005.

[5] M. Fazel, R. Ge, S. M. Kakade, and M. Mesbahi, “Global convergence of policy gradient methods for the linear quadratic regulator,” in Proc. 35th Int. Conf. Mach. Learn., 2018, pp. 10772–10782. [Online]. Available: http://arxiv.org/abs/1801.05039

[6] S. Bradtke, B. Ydstie, and A. Barto, “Adaptive linear quadratic control using policy iteration,” in Proc. Amer. Control Conf., 1999, pp. 3475–3479. [Online]. Available: https://ieeexplore.ieee.org/document/735224/
[7] F. L. Lewis and D. Vrabie, “Reinforcement learning and adaptive dynamic programming for feedback control,” IEEE Circuits Syst. Mag., vol. 9, no. 3, pp. 32–50, Third Quarter 2009. [Online]. Available: https://ieeexplore.ieee.org/document/5227780/

[8] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, “On the sample complexity of the linear quadratic regulator,” Found. Comput. Math., vol. 20, pp. 633–679, 2020. [Online]. Available: http://link.springer.com/10.1007/s10208-019-09426-y

[9] H. Mania, S. Tu, and B. Recht, “Certainty equivalence is efficient for linear quadratic control,” in Proc. Adv. Neural Inf. Process. Syst., 2019, Art. no. 911. [Online]. Available: https://proceedings.neurips.cc/paper/2019/file/5dc8e39f017e019430d5a162c3e3cb-Paper.pdf

[10] M. Abeille and A. Lazaric, “Efficient optimistic exploration in linear-quadratic regulators via Lagrangian relaxation,” in Proc. 37th Int. Conf. Mach. Learn., 2020, Art. no. 3.

[11] K. J. Astrom and B. Wittenmark, Adaptive Control, 2nd ed. Mineola, NY, USA: Dover, 2008.

[12] L. Ljung, System Identification: Theory for the User. Hoboken, NJ, USA: Prentice Hall, 1999.

[13] H. Rahimian and S. Mehrorta, “Frameworks and results in distributionally robust optimization,” Open J. Math. Optim., vol. 3, pp. 1–85, 2022. [Online]. Available: https://ojmo.centre-mersenne.org/articles/10.5802/ojmo.15.

[14] P. Coppens, M. Schuurmans, and P. Patrinos, “Data-driven distributionally robust LQR with multiplicative noise,” in Proc. 2nd Conf. Learn. Dyn. Control, 2020, pp. 521–530. [Online]. Available: http://proceedings.mlr.press/v120/coppens20a.html

[15] W. M. Wonham, “Optimal stationary control of a linear system with state-dependent noise,” SIAM J. Control, vol. 5, no. 3, pp. 486–500, 1967.

[16] T. Wang, H. Zhang, and Y. Luo, “Stochastic linear quadratic optimal control for model-free discrete-time systems based on Q-learning algorithm,” Neurocomputing, vol. 312, pp. 1–8, 2018. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0925231218304387

[17] B. Gravell, P. M. Esfahani, and T. Summers, “Learning robust control for LQR systems with multiplicative noise via policy gradient,” 2019. [Online]. Available: http://arxiv.org/abs/1905.13547

[18] B. Pang and Z.-P. Jiang, “Reinforcement learning for adaptive optimal stationary control of linear stochastic systems,” 2021. [Online]. Available: http://arxiv.org/abs/2107.07788

[19] T. Damm, Rational Matrix Equations in Stochastic Control, 1st ed. Berlin, Germany: Springer, 2004. [Online]. Available: http://link.springer.com/10.1007/b109096

[20] E. Todorov and M. Jordan, “Optimal feedback control as a theory of motor coordination,” Nature Neurosci., vol. 5, no. 11, pp. 1226–1235, 2002. [Online]. Available: http://www.nature.com/articles/nn963

[21] E. Todorov, “Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system,” Georgia Tech, vol. 17, no. 5, pp. 1084–1108, 2005. [Online]. Available: https://direct.mit.edu/neuroscience/article/175/1084-1108/6949

[22] G. Russo and R. Shorten, “On common noise-induced synchronization in complex networks with state-dependent noise diffusion processes,” Physica D, Nonlinear Phenomena, vol. 369, pp. 47–54, 2018. [Online]. Available: https://doi.org/10.1016/j.physd.2018.01.003

[23] R. R. Mohler and W. J. Kolodziej, “An overview of stochastic bilinear control processes,” IEEE Trans. Syst., Man, Cybern., vol. 10, no. 12, pp. 913–918, Dec. 1980. [Online]. Available: https://ieeexplore.ieee.org/document/4308421/

[24] R. Mohler, C. Bruni, and A. Gandolfi, “A systems approach to immunol coordination,” J. Franklin Inst., vol. 301, no. 1/2, pp. 103–122, Jan. 1976. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/0016087276901356

[25] A. Gershon and U. Shaked, “Multi-plicative noise and non-Gaussianity: A paradigm for atmospheric regimes,” J. Atmospheric Sci., vol. 62, no. 5, pp. 1391–1409, 2005.

[26] D. Bernstein, “Robust and optimal control for model-free discrete-time systems based on Q-learning algorithm,” IEEE Trans. Autom. Control, vol. AC-23, no. 4, pp. 913–918, Dec. 1987. [Online]. Available: https://ieeexplore.ieee.org/document/1104517/

[27] J. Doyle, “Guaranteed margins for LQR regulators,” IEEE Trans. Autom. Control, vol. AC-23, no. 4, pp. 756–757, Aug. 1978. [Online]. Available: http://ieeexplore.ieee.org/document/1101812/

[28] D. Blatt and L. El-Ghannouj, “State-feedback control of systems with multiplicative noise via linear matrix inequalities,” SIAM Control Optim., vol. 24, no. 3, pp. 223–239, 1995.

[29] P. Coppens and P. Patrinos, “Policy iteration using Q-functions: Linear dynamics with multiplicative noise,” 2022. arXiv:2212.01192

[30] C. S. Kubrusly, “Identification of discrete-stochastic bilinear systems,” IEEE Trans. Autom. Control, vol. AC-30, no. 11, pp. 1082–1087, Nov. 1985. [Online]. Available: https://ieeexplore.ieee.org/document/1103840/

[31] E. Todorov and M. Jordan, “Static L and H output-feedback of discrete-time LTI systems with state multiplicative noise,” Syst. Control Lett., vol. 55, no. 3, pp. 232–239, 2006.

[32] B. Gravell and T. Summers, “Robust learning-based control via bootstrapped multiplicative noise,” in Proc. 2nd Conf. Learn. Dyn. Control, 2020, pp. 599–607. [Online]. Available: https://proceedings.mlr.press/v120/gravell20a.html

[33] B. Di and A. Lamperski, “Confidence bounds on identification of linear systems with multiplicative noise,” in Proc. Amer. Control Conf., 2021, pp. 2212–2217.

[34] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” SIAM Rev., vol. 51, no. 3, pp. 455–500, 2009. [Online]. Available: http://epubs.siam.org/doi/10.1137/08071471X

[35] B. Gravell and T. Summers, “Robust learning-based control via bootstrapped multiplicative noise,” in Proc. 2nd Conf. Learn. Dyn. Control, 2020, pp. 599–607. [Online]. Available: https://proceedings.mlr.press/v120/gravell20a.html

[36] B. Di and A. Lamperski, “Confidence bounds on identification of linear systems with multiplicative noise,” in Proc. Amer. Control Conf., 2021, pp. 2212–2217.

[37] T. G. Kolda, “Multilinear operators for higher-order decompositions,” Sandia National Laboratories (SNL), Tech. Rep. SAND2006-2081, 2006. [Online]. Available: http://wwwosti.gov/servlets/purl/923081-u0xXJa/

[38] T. G. Kolda and B. W. Bader, “Tensor decompositions and applications,” SIAM Rev., vol. 51, no. 3, pp. 455–500, 2009. [Online]. Available: http://epubs.siam.org/doi/10.1137/08071471X

[39] B. Hua and I. Ehsani, “Positive linear maps and the Lyapunov equation,” in Proc. Linear Operators Matrices, 2002, pp. 107–120.

[40] T. Damm, “Stability of linear systems and positive semigroups of symmetric matrices,” in Proc. Posit. Syst., 2003, pp. 207–214.

[41] G. Lindblad, “On the generators of quantum dynamical semigroups,” Commun. Math. Phys., vol. 48, no. 2, pp. 119–130, 1976. [Online]. Available: http://link.springer.com/10.1007/BF01680499

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
[51] D. Romero and G. Leus, “Wideband spectrum sensing from compressed measurements using spectral prior information,” IEEE Trans. Signal Process., vol. 61, no. 24, pp. 6232–6246, Dec. 2013. [Online]. Available: https://ieeexplore.ieee.org/document/6609101/

[52] A. Kukush, I. Markovsky, and S. V. Huffel, “Consistent estimation in the bilinear multivariate errors-in-variables model,” Metrika, vol. 57, no. 3, pp. 253–285, 2003. [Online]. Available: http://link.springer.com/10.1007/s001840200217

[53] J. A. Tropp, “An introduction to matrix concentration inequalities,” Found. Trends Mach. Learn., vol. 8, no. 1/2, pp. 1–230, 2015. [Online]. Available: http://dx.doi.org/10.1561/22000000048

[54] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, “Robustness,” in Handbook of Semidefinite Programming. Berlin, Germany: Springer, 2000. [Online]. Available: http://link.springer.com/10.1007/978-1-4615-4381-7_6

[55] P. Coppens and P. Patrinos, “Sample complexity of data-driven stochastic LQR with multiplicative uncertainty,” in Proc. IEEE 59th Conf. Decis. Control, 2020, pp. 6210–6215.

[56] P. Coppens and P. Patrinos, “Provably stable learning control of linear dynamics with multiplicative noise,” 2023, arXiv:2207.06062.

[57] P. Coppens and P. Patrinos, “Data-driven distributionally robust MPC for constrained stochastic systems,” IEEE Control Syst. Lett., vol. 6, pp. 1274–1279, 2022.

[58] R. A. Horn and C. R. Johnson, Topics in Matrix Analysis. Cambridge, U.K.: Cambridge Univ. Press, 1991. [Online]. Available: https://www.cambridge.org/core/product/identifier/9780521430538/type/book

[59] E. de Klerk, Aspects of Semidefinite Programming. Berlin, Germany: Springer, 2002. [Online]. Available: http://link.springer.com/10.1007/b105286

[60] M.-D. Choi, “Positive linear maps on complex matrices,” http://dx.doi.org/10.1007/BF01214346

[61] R. A. Horn and C. R. Johnson, Matrix Analysis, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2012. [Online]. Available: https://dl.acm.org/doi/book/10.5555/2422911

[62] R. Jungers, The Joint Spectral Radius, vol. 385. Berlin, Germany: Springer, 2009. [Online]. Available: http://link.springer.com/10.1007/978-3-540-95980-9

[63] G. H. Golub and C. F. V. Loan, Matrix Computations, 4th ed.. Baltimore, MD, USA: JHU Press, 2013.

[64] N. Lee and A. Cichocki, “Fundamental tensor operations for large-scale applications.”

[65] G. R. Grimmett and D. Stirzaker, Probability and Random Processes. Oxford, U.K.: Oxford Univ. Press, 2001.

[66] V. Balakrishnan and L. Vandenberghe, “Semidefinite programming duality and linear time-invariant systems,” IEEE Trans. Autom. Control, vol. 48, no. 1, pp. 30–41, Jan. 2003.

[67] T. Morozan, “Stabilization of some stochastic discrete-time control systems,” Stochastic Anal. Appl., vol. 1, no. 1, pp. 89–116, 1983.

[68] D. Testa, “3 × 3 singular matrices of linear forms,” 2017, arXiv:1701.06864.

[69] L. Lovász, “Singular spaces of matrices and their application in combinatorics,” Bol. Soc. Bras. Mat., vol. 20, no. 1, pp. 87–99, 1989.

[70] B. Chen and Z. Li, “On the tensor spectral p-norm and its dual norm via partitions,” Comput. Optim. Appl., vol. 75, no. 3, pp. 609–628, 2020. [Online]. Available: https://doi.org/10.1007/s10589-020-00177-z

[71] M. Simchowitz and D. Foster, “Naive exploration is optimal for online LQR,” in Proc. 37th Int. Conf. Mach. Learn., 2020, pp. 8937–8948. [Online]. Available: https://proceedings.mlr.press/v119/simchowitz20a.html

[72] E. Delage and Y. Ye, “Distributionally robust optimization under moment uncertainty with application to data-driven problems,” Operations Res., vol. 58, no. 3, pp. 595–612, 2010. [Online]. Available: http://pubsonline.informs.org/doi/abs/10.1287/opre.1090.0741

[73] S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[74] M. Ledoux and M. Talagrand, Probability in Banach Spaces. Berlin, Germany: Springer, 1991. [Online]. Available: http://link.springer.com/10.1007/978-3-642-20212-4

[75] S. Boucheron, G. Lugosi, and P. Massart, Concentration Inequalities. Oxford, U.K.: Oxford Univ. Press, 2013. [Online]. Available: https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199535255.001.0001/acprof-9780199535255

[76] T. N. E. Greville, “Note on the generalized inverse of a matrix product,” SIAM Rev., vol. 8, no. 4, pp. 518–521, 1966. [Online]. Available: http://epubs.siam.org/doi/10.1137/1008107

Peter Coppens received the M.Eng. degree in mathematical engineering from the KU Leuven, Leuven, Belgium, in 2019. He is currently a Ph.D. Researcher with the Department of Electrical Engineering (ESAT), KU Leuven. His current research interests include learning control, enabled by distributionally robust optimization, specifically its use in model predictive control and safety critical applications.

Mr. Coppens was the recipient of an FWO fundamental research fellowship.

Panagiotis (Panos) Patrinos received the Ph.D. degree in control and optimization, the M.S. degree in applied mathematics, and the M.Eng. degree in chemical engineering from the National Technical University of Athens, Athens, Greece, in 2010, 2005, and 2003, respectively. He is currently an Associate Professor with the Department of Electrical Engineering (ESAT), KU Leuven, Leuven, Belgium. In 2014, he was a Visiting Professor with Stanford University. After his Ph.D., he held Postdoc positions with the University of Trento and IMT Uuccia, Italy, where he became an Assistant Professor in 2012. His current research interests include the intersection of optimization, control, and learning. In particular, he is interested in the theory and algorithms for structured nonconvex optimization as well as learning-based, model predictive control with a wide range of applications including autonomous vehicles, machine learning, and signal processing.

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.