2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Stability Bounds for Learning-Based Adaptive Control of Discrete-Time Multi-Dimensional Stochastic Linear Systems with Input Constraints

Seth Siriya, Jingge Zhu, Dragan Nešić, and Ye Pu

Abstract— **We consider the problem of adaptive stabilization for discrete-time, multi-dimensional linear systems with bounded control input constraints and unbounded stochastic disturbances, where the parameters of the system are unknown. To address this challenge, we propose a certainty-equivalent control scheme combining online parameter estimation with saturated linear control. We establish the existence of a high probability stability bound on the closed-loop system, under additional assumptions on the system and noise processes. Numerical examples are presented to illustrate our results.**

I. Introduction

Adaptive control (AC) is concerned with the design of controllers for dynamical systems whose model parameters are unknown. When deploying these algorithms in the real world, it is important that actuator saturation is accounted for during design, since ignorance of such issues may result in failure to achieve stability. Moreover, systems are sometimes subject to rare, large, disturbances — often modelled by additive, unbounded stochastic noise — which can degrade control performance and potentially cause instability. This motivates the need to develop provably stabilizing adaptive control algorithms that simultaneously handle input constraints and additive, unbounded, stochastic disturbances.

Discrete-time (DT) stochastic AC has recently seen interest in the form of online model-based reinforcement learning — especially for the online linear quadratic regulation (LQR) task, which aims to minimize regret with respect to the optimal LQR controller on an unknown, linear system with additive, sub-Gaussian disturbances (see [1], [2]). These results have been extended to handle state and input constraints [3], but only when disturbances are bounded. DT extremum seeking (ES) AC results have also shown promise for stabilizing unstable DT systems ([4], [5]), but do not account for input constraints. Despite the history of DT stochastic AC, tracing back to classic linear results such as [6], [7], the control of non-strictly stable systems subject to hard input constraints has not garnered attention. Other nonlinear DT stochastic AC problems have been considered, such as deadzone nonlinearities [8], and linearly parameterized nonlinear systems [9]. On the other hand, control constraints have been studied for the stabilization of unknown, DT output-feedback linear systems with bounded disturbances in [10], [11], but unbounded disturbances are not supported. Recently, mean square boundedness of a learning-based adaptive control scheme for at-worst marginally stable, scalar, linear systems, with Gaussian disturbances and bounded controls, was established in [12], by combining results from non-asymptotic system identification [13] with input-constrained stochastic control [14]. However, multi-dimensional results are missing.

Motivated by our previous scalar result [12], we move towards filling the gap in the multi-dimensional setting. In particular, we aim to develop a method for adaptive stabilization of unknown, multi-dimensional linear systems, subject to additive, i.i.d. sub-Gaussian zero-mean stochastic disturbances, and hard, positive upper bound constraints on the control magnitude. Our main contributions are twofold.

Firstly, we propose a certainty-equivalence (CE) adaptive control scheme to address the problem. It consists of a saturated linear controller based on parameter estimates obtained via ordinary least squares (OLS) online, which has been intentionally excited by a bounded noise to facilitate parameter convergence. The saturation level and exciting noise level can be jointly selected to satisfy the control input constraint. Moreover, we do not assume prior knowledge of any bounds on the system parameters.

Secondly, we prove the existence of a high probability stability bound which holds on sub-sampled states of the closed-loop system, under the assumption that the system is controllable, $\|A\| \leq 1$, the saturation level of the controller is sufficiently large for the given disturbance and exciting noise processes, and that a persistency of excitation-like condition holds on the state-input data sequence. To achieve this, we first establish an upper bound on the parameter estimation error that holds over time with high probability using recent results on non-asymptotic system identification [13]. Then, we derive a probabilistic upper bound on the norm of the sub-sampled states which relies on a given estimation error bound. These two results are subsequently combined to derive a parameterized family of high probability upper bounds on the norm of the sub-sampled states. Our main result then follows.

Notation: For $x \in \mathbb{R}^n$, $|x|$ denotes its Euclidean norm. For any $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, (x, y) stands for $\begin{bmatrix} x^\top & y^\top \end{bmatrix}^\top$. Given a matrix $M \in \mathbb{R}^{n \times m}$, $\|M\|$ is its induced 2-norm, $\sigma_{\max}(M)$ and $\sigma_{\min}(M)$ denotes its maximum and minimum singular values respectively, $\mathcal{B}_r(M)$ denotes the 2-norm open ball of radius $r > 0$ centered at M and $\bar{\mathcal{B}}_r(M)$ denotes its closure, and $M^\dagger$ denotes its Moore-Penrose inverse. Given $M \in \mathbb{R}^{n \times n}$, $\rho(M)$ denotes its spectral radius, and if it is

The work of Seth Siriya was supported by the Australian Government Research Training Program (RTP) Scholarship. The work of Dragan Nešić was supported by the Australian Research Council through the Discovery Project under Grant DP210102600. The work of Ye Pu was supported by the Australian Research Council under Project DE220101527.

The authors are with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, VIC 3010, Australia (e-mail: ssiriya@student.unimelb.edu.au, {jingge.zhu, dnesic, ye.pu}@unimelb.edu.au).

symmetric, $\lambda_{\min}(M)$ denotes its minimum eigenvalue, and $\lambda_{\max}(M)$ denotes its maximum eigenvalue. For $r > 0$, we define the saturation function $\mathrm{sat}_r : \mathbb{R}^d \to \mathbb{R}^d$ by $\mathrm{sat}_r(x) := x$ if $|x| \le r$, and $\mathrm{sat}_r(x) := rx/|x|$ if $|x| > r$. For any $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$, define $\mathcal{R}_\kappa(A,B) := \begin{bmatrix} B & AB & \dots & A^{\kappa-1}B \end{bmatrix}$ as the corresponding κ-step reachability matrix. The identity matrix is denoted by $\mathbf{I}$. Denote the unit sphere embedded in $\mathbb{R}^d$ by $\mathcal{S}^{d-1}$. Given sets A and B, A^{c} denotes the complement of A, $A \cap B$ denotes the intersection of A and B, and $A \cup B$ denotes their union. Consider a probability space $(\Omega, \mathcal{F}, P)$, and a random vector $X : \Omega \to \mathbb{R}^d$, an event $A \in \mathcal{F}$, and a sub-sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$, defined on this space. The expected value of X is denoted by $\mathbb{E}[X]$. We define the indicator function $\mathbf{1}_A : \Omega \to \{0,1\}$ as $\mathbf{1}_A := 1$ on the event A, and $\mathbf{1}_A := 0$ on the event A^{c}. For scalar X, we say X is σ^2-sub-Gaussian if $\mathbb{E}[e^{tX}] \le e^{\sigma^2 t^2/2}$ for all $t \in \mathbb{R}$. For vector X, we say X is σ^2-sub-Gaussian if $\mathbb{E}[X] = 0$, and $u^\top X$ is σ^2-sub-Gaussian for any $u \in \mathcal{S}^{d-1}$.

II. Problem Setup

Consider the following stochastic linear system:

$$X_{t+1} = AX_t + BU_t + W_t,\ t \in \mathbb{N}_0, \quad X_0 = x_0, \qquad (1)$$

where the random sequences $(X_t)_{t\in\mathbb{N}_0}$, $(U_t)_{t\in\mathbb{N}_0}$ and $(W_t)_{t\in\mathbb{N}_0}$ are the states, controls, and disturbances, taking values in $\mathbb{R}^n$, $\mathbb{R}^m$, and $\mathbb{R}^n$ respectively, $x_0 \in \mathbb{R}^n$ is the initial state, and $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are the true, unknown, system matrices. For convenience, $\theta_* = \begin{bmatrix} A & B \end{bmatrix}$ denotes the true system parameter. This is in contrast to $\hat{\theta}_t$, denoting the estimated parameter at time t, and will be formally defined later. We make the following assumptions on (1).

Assumption 1. *The disturbance $(W_t)_{t\in\mathbb{N}_0}$ is an i.i.d. sequence that has an unbounded support, is mean-zero and σ_W^2-sub-Gaussian, with a covariance matrix Σ_W.*

Assumption 2. *The matrix A is full rank, satisfies $\|A\| \le 1$, and (A,B) is κ-step reachable with $\kappa \le n$, that is, $\mathrm{rank}(\mathcal{R}_\kappa(A,B)) = n$. For ease of notation, we denote $\mathcal{R}_* = \mathcal{R}_\kappa(A,B)$.*

Remark 1. Note that Assumption 1 is broad enough to handle many different types of disturbance with an unbounded support, including Gaussian distributions. We assume i.i.d. sequences to simplify the exposition, but our analysis can be extended to martingale difference sequences. Full rank A is required in Assumption 2 to establish the perturbation bounds for our saturated linear controller introduced later. It is satisfied in many systems, such as linear systems discretized with zero-order hold. Moreover, we assume $\|A\| \le 1$ and (A,B) reachable, for which in the non-adaptive case there exist bounded control policies rendering such systems mean square bounded with unbounded stochastic disturbances [15]. This gives us hope that stabilizing adaptive control is possible. In contrast, when $\rho(A) > 1$, [14] showed that mean square boundedness is impossible with unbounded disturbances and bounded controls. Although the plants we consider are marginally stable in the deterministic setting and hence bounded, they exhibit unbounded behaviour in the stochastic setting. For example, the scalar system $X_{t+1} = X_t + W_t$ with $X_0 = 0$ and $W_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ has variance $\mathrm{Var}\,(X_t) = t$, which exhibits unbounded growth.

Our goal is to formulate an adaptive control policy $(\pi_t)_{t\in\mathbb{N}_0}$ such that π_t is a mapping from current and past state and control input data $(X_0, \dots, X_t, U_0, \dots, U_{t-1})$ and a randomizaton term V_t to $\mathbb{R}^m$ for $t \in \mathbb{N}_0$. Here, $(V_t)_{t\in\mathbb{N}_0}$ taking values in $\mathbb{R}^m$ is an i.i.d. random sequence whose purpose is to excite the system in order to facilitate convergence of parameter estimates. The overall policy needs to be designed so that $|U_t| \le U_{\max}$ holds where $U_{\max} > 0$ is the control magnitude constraint, whilst provably achieving stochastic stability guarantees on the closed-loop system states $(X_t)_{t\in\mathbb{N}}$ with $U_t = \pi_t(X_0, \dots, X_t, U_0, \dots, U_{t-1}, V_t)$. Moreover, we require in our design that π_t does not depend on θ_*.

III. Method and Main Result

For controller design, we require knowledge of some κ satisfying Assumption 2. Although we can have $\kappa < n$ in many cases when systems have multiple inputs, if it is only known that (A,B) is controllable, $\kappa = n$ is always a valid choice. Our control strategy is summarized in Algorithm 1.

Algorithm 1 Stochastic Adaptive Input-Constrained Control

> **Inputs:** $U_{\max} > 0$, $C \in (0, U_{\max})$, $\bar{A}_0 \in \mathbb{R}^{n\times n}$, $\bar{B}_0 \in \mathbb{R}^{n\times m}$, $\kappa \le n$
> $D \leftarrow U_{\max} - C$
> $\tau \leftarrow 0$
> Measure X_0
> **for** $\tau = 0, 1, \dots$ **do**
>   Sample $V_{\kappa\tau}, \dots, V_{\kappa(\tau+1)-1}$ satisfying Assumption 3
>   Compute $\bar{U}_\tau$ following (2)
>   **for** $t = \kappa\tau, \dots, \kappa(\tau+1)-1$ **do**
>   Extract U_t from $\bar{U}_\tau$ following (2)
>   Apply U_t to (1)
>   Measure X_{t+1} from (1)
>   **end for**
>   Compute $\bar{\theta}_{\tau+1}$ following (3)
> **end for**

We now describe our strategy in greater detail. For all $\tau \in \mathbb{N}_0$, our sub-sampled control sequence $(\bar{U}_\tau)_{\tau\in\mathbb{N}_0}$ is

$$\underbrace{\begin{bmatrix} U_{\kappa(\tau+1)-1} \\ \vdots \\ U_{\kappa\tau} \end{bmatrix}}_{\bar{U}_\tau} = \mathrm{sat}_D(-g(\bar{A}_\tau, \bar{B}_\tau)X_{\kappa\tau}) + \begin{bmatrix} V_{\kappa(\tau+1)-1} \\ \vdots \\ V_{\kappa\tau} \end{bmatrix} \qquad (2)$$

where $g(\bar{A}, \bar{B}) := \mathcal{R}_\kappa(\bar{A}, \bar{B})^\dagger(\bar{A})^\kappa$ for $\begin{bmatrix} \bar{A} & \bar{B} \end{bmatrix} \in \mathbb{R}^{n\times(n+m)}$, and V_t is an additive *excitation term* sampled so that Assumption 3 is satisfied.

Remark 2. When the true system parameter is used for control, such that $\mathrm{sat}_D(-g(A,B)x)$ is our control law, the policy can be viewed as a saturated deadbeat controller for the dynamical system obtained when (1) is sampled with periodicity κ. A similar controller structure was shown to achieve mean square boundedness in [15], except the saturation and and linear gain is switched. We opt for our order since our CE control strategy involves using estimates $(\bar{A}_\tau, \bar{B}_\tau)$ rather than (A, B), and our estimates can be unbounded leading to unbounded gain. Applying saturation afterwards guarantees our controls satisfy $U_{\max}$.

Assumption 3. *The random sequence $(V_t)_{t\in\mathbb{N}_0}$ taking values in $\mathbb{R}^m$ is i.i.d. Additionally, $|V_t| \le C$ holds, and V_i,W_j are independent for all $i, j \in \mathbb{N}_0$.*

Remark 3. Assumption 3 restricts the magnitude of the additive noise V_t, which is required for satisfying control input constraints. We denote its covariance matrix by Σ_V.

Let $(\hat{\theta}_t)_{t\in\mathbb{N}}$ taking values in $\mathbb{R}^{n\times(n+m)}$ be the sequence of estimates of θ_* via OLS estimation:

$$\hat{\theta}_t \in \arg\min_{\theta\in\mathbb{R}^{n\times(n+m)}} \sum_{s=1}^{t} \|X_s - \theta Z_s\|_2^2\,, \tag{3}$$

where $(Z_t)_{t\in\mathbb{N}}$ taking values in $\mathbb{R}^{n+m}$ is the state-input data sequence, i.e. $Z_t = (X_{t-1}, U_{t-1})$. Let $(\bar{\theta}_\tau)_{\tau\in\mathbb{N}}$ be the sequence of sub-sampled parameter estimates satisfying $\bar{\theta}_\tau = \hat{\theta}_{\kappa\tau}$, and let $(\bar{A}_\tau)_{\tau\in\mathbb{N}}, (\bar{B}_\tau)_{\tau\in\mathbb{N}}$, be the sub-sampled estimates of A and B respectively, satisfying $\begin{bmatrix}\bar{A}_\tau & \bar{B}_\tau\end{bmatrix} = \bar{\theta}_\tau$. Note, the initial parameter estimate $(\bar{A}_0, \bar{B}_0)$ is not computed via OLS, but instead freely chosen by the designer in $\mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$. Additionally, C is a user-specified excitation constant satisfying $0 < C < U_{\max}$ which determines the size of the excitation term, and $D = U_{\max} - C$ is the magnitude of the certainty-equivalent component of the control policy.

Under this control strategy, the sub-sampled state sequence $(\bar{X}_\tau)_{\tau\in\mathbb{N}_0}$, satisfying $\bar{X}_\tau = X_{\kappa\tau}$ evolves via

$$\begin{aligned}\bar{X}_{\tau+1} &= AX_{\kappa(\tau+1)-1} + BU_{\kappa(\tau+1)-1} + W_{\kappa(\tau+1)-1}\\ &= A^\kappa \bar{X}_\tau + \mathcal{R}_*\mathrm{sat}_D(-g(\bar{A}_\tau, \bar{B}_\tau)\bar{X}_\tau) + \bar{V}_\tau + \bar{W}_\tau\end{aligned} \tag{4}$$

for all $\tau \in \mathbb{N}_0$, where $\bar{V}_\tau = \mathcal{R}_*[V_{\kappa(\tau+1)-1}^\top, \ldots, V_{\kappa\tau}^\top]^\top$, and $\bar{W}_\tau = \mathcal{R}_\kappa(A, I)[W_{\kappa(\tau+1)-1}^\top, \ldots, W_\tau^\top]^\top$. Next, we let $M_{\bar{W}} = \ln(\mathbb{E}[e^{|\bar{W}_\tau|}])$, and $M_{\bar{V}} = \ln(\mathbb{E}[e^{|\bar{V}_\tau|}])$, whose existence are guaranteed from Assumptions 1 and 3. We assume they relate to D and $\mathcal{R}_*$ as follows.

Assumption 4. *The saturation level D, $M_{\bar{V}}$, and $M_{\bar{W}}$, satisfy $\frac{D}{\|\mathcal{R}_*^\dagger\|} > M_{\bar{V}} + M_{\bar{W}}$.*

Remark 4. The term $\frac{D}{\|\mathcal{R}_*^\dagger\|}$ in Assumption 4 can be interpreted as the influence of the control strategy on the system. Conversely, $M_{\bar{V}}$ and $M_{\bar{W}}$ are statistics that can be viewed as the influence of the exciting noise and disturbance on the system. Roughly, they characterize the variance of the noise. Assumption 4 ensures the controls can overpower the disturbance/noise stochastically, allowing us to invoke Lyapunov-type arguments for stability analysis. For any D, it is satisfied so long as $M_{\bar{W}}$ and $M_{\bar{V}}$ are sufficiently small.

We define the block martingale small-ball (BMSB) condition, and assume the state-input data sequence satisfies it.

Definition 1. (Martingale Small-Ball [13, Definition 2.1]) Given process $(Z_t)_{t\ge1}$ taking values in $\mathbb{R}^d$, we say it satisfies the $(k, \Gamma_{\text{sb}}, p)$-block martingale small-ball (BMSB) condition for $k \in \mathbb{N}$, $\Gamma_{\text{sb}} \succ 0$, and $p > 0$, if, for any $\zeta \in \mathcal{S}^{d-1}$ and $j \ge 0$, $\frac{1}{k}\sum_{i=1}^{k} P(|\zeta^\top Z_{j+i}| \ge \sqrt{\zeta^\top\Gamma_{\text{sb}}\zeta} \mid \mathcal{F}_j) \ge p$ holds. $(\mathcal{F}_t)_{t\ge1}$ is any filtration which $(\zeta^\top Z_t)_{t\ge1}$ is adapted to.

Assumption 5. *The constants $k > 0$, $\Gamma_{sb} \succ 0$, and $p > 0$ are such that the state-input data sequence $(Z_t)_{t\in\mathbb{N}}$ satisfies the (k, Γ_{sb}, p)-BMSB condition.*

Remark 5. The BMSB condition in Definition 1 can be used to establish that persistency of excitation holds, which is important for deriving high probability bounds on the estimation error (see [16]). By supposing $(Z_t)_{t\in\mathbb{N}}$ satisfies the BMSB condition in Assumption 5, conditioned on past Z_t, the averaged distributions of future Z_t are sufficiently spread. We proved this holds in the scalar case [12] using exciting noise, and leave the vector case to future work.

We now present the main result on the existence of a high probability stability bound for our control scheme.

Theorem 1. *Suppose Assumptions 1-5 hold. There exist $\lambda \in (0, 1)$, $N_1, N_3 > 0$, and strictly increasing function $N_2 : \mathbb{R}_{\ge0} \to \mathbb{R}_{\ge0}$ such that*

$$|\bar{X}_\tau| \le \ln(N_2(|x_0|)(2/\delta)^{N_1}\lambda^\tau + N_3) + \ln(2/\delta) \tag{5}$$

holds with probability at least $1 - \delta$ for all $x_0 \in \mathbb{R}^n$, $\delta \in (0, 1)$ and $\tau \in \mathbb{N}_0$.

Theorem 1 says that, for any initial state $x_0 \in \mathbb{R}^n$, sub-sampled time $\tau \in \mathbb{N}_0$, and failure probability $\delta \in (0, 1)$, with probability at least $1 - \delta$, $\bar{X}_\tau$ will be in a ball around the origin with size equal to the right hand side of (5). In particular, we can interpret the result as a stability bound since the right hand side is uniformly bounded by $\ln(N_2(|x_0|)(2/\delta)^{N_1} + N_3) + \ln(2/\delta)$ over all $\tau \in \mathbb{N}_0$, and will asymptotically converge to $\ln(N_3) + \ln(2/\delta)$ regardless of x_0. Although the structure of this bound is non-standard, it can show up when bounding systems which converge to a set linearly, rather than exponentially. Moreover, the right hand side increases as δ approaches zero. This reflects the fact that more conservative bounds can assert statements with greater confidence, and is a common phenomenon in the statistical learning literature.

IV. Proof Sketch of Main Result

We now build towards the proof of Theorem 1. Firstly, in Section IV-A, we establish a high probability upper bound on the parameter estimation error in the form of Proposition 1. Next, in Section IV-B, we provide Lemmas 1 and 2, which are used to obtain Proposition 2 — a probabilistic upper bound on the norm of the sub-sampled states which relies on a given estimation error bound. This result is subsequently combined with Proposition 1 to derive a family of high probability stability bounds in Theorem 2. Theorem 1 then follows after simplification. Although we provide sketches of the key ideas behind these results here, we defer the formal proofs to the supplementary materials on Arxiv [17].

A. Estimation Error Bound

Under Assumptions 1, 2, 3, and 5, we provide Proposition 1 — a high probability error bound on our parameter estimates from (3). In particular, it says that, with probability at least $1-\delta$, the function $e(T,\delta,x_0)$ will bound the estimation error $\|\hat{\theta}_T-\theta_*\|$ uniformly over *all* T greater than $T_0(\delta,x_0)$. We derive this result by making use of [13, Theorem 2.4], which provides high probability upper bounds on the estimation error for parameter estimates obtained by applying OLS to a general time-series with linear responses, and can successfully be applied when the BMSB condition in Definition 1 is satisfied, and high-probability upper bounds on $\sum_{t=1}^{T} Z_t Z_t^\top$ can be found. One key difference between Proposition 1 and [13, Theorem 2.4] is the latter holds at any sufficiently large T with high probability, rather than with high probability over all $T \geq T_0(\delta,x_0)$.

Proposition 1. *Suppose Assumptions 1, 2, 3, and 5 hold. Consider the sequence of parameter estimates $(\hat{\theta}_t)_{t\in\mathbb{N}}$ from (3). Fix $\delta\in(0,1)$, and $x_0\in\mathbb{R}^n$. Then,*

$$\|\hat{\theta}_T-\theta_*\|_2 \leq e(T,\delta,x_0) \text{ for all } T\geq T_0(\delta,x_0)$$

holds with probability at least $1-\delta$, where

$$\begin{aligned} e(T,\delta,x_0) :=\ & \frac{90\sigma_W}{p}\Big((T\lambda_{min}(\Gamma_{sb}))^{-1}\Big(n \\ & + (n+m)\ln(10/p)) + \ln\det\Big(\frac{\pi^2T^2}{2\delta} \\ & \times (4|x_0|^2 + 2(D^2+\operatorname{tr}(\Sigma_V)) + 4(\|B\|^2(D^2+\operatorname{tr}(\Sigma_V)) \\ & + \operatorname{tr}(\Sigma_W))T^2 + \lambda_{max}(\Gamma_{sb}))\Gamma_{sb}^{-1}\Big) + \ln\Big(\frac{\pi^2T^2}{2\delta}\Big)\Big)\Big)^{1/2}, \end{aligned} \tag{6}$$

$$\begin{aligned} & T_0(\delta,x_0) \\ & := \min\{T_0'\in\mathbb{N} \mid T\geq \frac{10k}{p^2}\Big(2(n+m)\ln(10/p) \\ & + \ln\det\Big(\frac{\pi^2T^2}{2\delta}(4|x_0|^2+2(D^2+\operatorname{tr}(\Sigma_V)) \\ & + 4(\|B\|^2(D^2+\operatorname{tr}(\Sigma_V)) + \operatorname{tr}(\Sigma_W))T^2 + \\ & \lambda_{max}(\Gamma_{sb}))\Gamma_{sb}^{-1}\Big) + \ln\Big(\frac{\pi^2T^2}{2\delta}\Big)\Big) \text{ for all } T\geq T_0'\}. \end{aligned} \tag{7}$$

We describe the key ideas for proving Proposition 1. Firstly, we establish that an estimation error bound holds at any specific time T with probability $1-\delta'$ (where δ' is the failure probability) by verifying the conditions for [13, Theorem 2.4], treating $Z_t \leftarrow (X_{t-1},U_{t-1})$ as the covariates and $Y_t \leftarrow X_t$ as the response. BMSB holds by Assumption 5, and $P(\sum_{t=1}^{T} Z_t Z_t^\top \npreceq T\overline{\Gamma}) \leq \delta'$ is established with $\overline{\Gamma} \leftarrow (1/\delta')(4|x_0|^2 + 2(D^2+\operatorname{tr}(\Sigma_V)) + 4(\|B\|^2(D^2+\operatorname{tr}(\Sigma_V)) + \operatorname{tr}(\Sigma_W))T^2 + \lambda_{\max}(\Gamma_{\mathrm{sb}}))$. We move from the specific T bound, to Proposition 1 that holds with probability $1-\delta$ (where δ is the failure probability) uniformly over all sufficiently large T, by replacing $\delta' \leftarrow 2\delta/(\pi^2T^2)$, and invoking a union bound argument.

B. Stability Bound

We state Lemma 1 on perturbation bounds for the CE component of the controls as a function of parameter estimation error and saturation level D.

Lemma 1. *Suppose Assumption 2 holds. There exist $m_q, M_q>0$ such that for all $D>0$ and $\epsilon\in[0,m_q]$,*

$$|\operatorname{sat}_D(-g(\bar{A},\bar{B})x) - \operatorname{sat}_D(-g(A,B)x)| \leq M_q\cdot D\cdot\epsilon$$

holds for all $\bar{A}\in\bar{\mathcal{B}}_\epsilon(A)$, $\bar{B}\in\bar{\mathcal{B}}_\epsilon(B)$, and $x\in\mathbb{R}^n$.

The result in Lemma 1 follows from a perturbation bound we derive on the controller saturation error using matrix analysis, which is convex in ϵ over a half-open interval. Of note in the proof, is the perturbation bound holds uniformly over all states. Intuitively, this is because after fixing D and ϵ sufficiently small, we can find a compact set S of x such that on S^{c}, both $\operatorname{sat}_D(-g(\bar{A},\bar{B})x)$ and $\operatorname{sat}_D(-g(A,B)x)$ are saturated, so the CE error will not grow with $|x|$. Although the CE error will grow with $|x|$ within S, we obtain a uniform bound based on the worst $|x|$.

We make use of Lemma 1 to derive Lemma 2. It bounds the expected value of an exponential function of the states of a family of systems that evolve similar to (4), but are instead parameterized by deterministic estimates of θ_* that are used for the controller. This bound holds uniformly over estimates in a sufficiently small ball around θ_*.

Lemma 2. *Suppose Assumptions 1, 2, and 3 hold. Let $m_q, M_q>0$ satisfy Lemma 1. Fix $\epsilon\in[0,m_q]$. Consider a family of random sequences $(Z_\tau^{\bar{A},\bar{B}})_{\tau\in\mathbb{N}_0}$ parameterized by $\bar{A}\in\bar{\mathcal{B}}_\epsilon(A), \bar{B}\in\bar{\mathcal{B}}_\epsilon(B)$ that evolve according to the closed-loop system:*

$$\begin{aligned} Z_{\tau+1}^{\bar{A},\bar{B}} =\ & A^\kappa Z_\tau^{\bar{A},\bar{B}} + \mathcal{R}_* sat_D(-\mathcal{R}_\kappa(\bar{A},\bar{B})^\dagger \bar{A}^\kappa Z_\tau^{\bar{A},\bar{B}}) \\ & + \tilde{V}_\tau + \tilde{W}_\tau, \end{aligned}$$

for $\tau\in\mathbb{N}_0$, where:

1) *$(\tilde{W}_\tau)_{\tau\in\mathbb{N}_0}$ and $(\tilde{V}_\tau)_{\tau\in\mathbb{N}_0}$ are i.i.d. random sequences with the same distribution as $(\bar{W})_{\tau\in\mathbb{N}_0}$ and $(\bar{V}_\tau)_{\tau\in\mathbb{N}_0}$ respectively;*
2) *$Z_0^{\bar{A},\bar{B}} = \underline{Z}$ for all $\bar{A}\in\bar{\mathcal{B}}_\epsilon(A)$ and $\bar{B}\in\bar{\mathcal{B}}_\epsilon(B)$, where $\underline{Z}$ is a random variable with distribution $\mu(\cdot)$;*
3) *$(\tilde{W}_\tau)_{\tau\in\mathbb{N}_0}$, $(\tilde{V}_\tau)_{\tau\in\mathbb{N}_0}$, $\underline{Z}$ are all independent.*

Then,

$$\mathbb{E}[e^{|Z_\tau^{\bar{A},\bar{B}}|} \mid Z_0^{\bar{A},\bar{B}} = z] \leq \quad \lambda^\tau(\epsilon)e^{|z|} + \beta(\epsilon)\sum_{k=0}^{\tau-1}\lambda^{\tau-1-k}(\epsilon) \tag{8}$$

for all $\bar{A}\in\bar{\mathcal{B}}_\epsilon(A)$, $\bar{B}\in\bar{\mathcal{B}}_\epsilon(B)$, $\tau\in\mathbb{N}_0$, and z in the support of $\mu(\cdot)$, where

$$\lambda(\epsilon) := e^{\|\mathcal{R}_*\|\cdot M_q\cdot D\cdot\epsilon + \frac{-D}{\|\mathcal{R}_*^\dagger\|} + M_{\bar{V}} + M_{\bar{W}}}, \tag{9}$$

$$\beta(\epsilon) := e^{\|\mathcal{R}_*\|\cdot M_q\cdot D\cdot\epsilon + M_{\bar{V}} + M_{\bar{W}}}. \tag{10}$$

We describe the key ideas for proving Lemma 2. Assuming $\bar{A}\in\bar{\mathcal{B}}_\epsilon(A)$ and $\bar{B}\in\bar{\mathcal{B}}_\epsilon(B)$, we can bound $|\bar{Z}_1^{\bar{A},\bar{B}}|$ based on the dynamics of the nominal sub-sampled closed-loop system (where the true parameter A,B is used for the controller), and the CE error, which is upper bounded using Lemma 1. This implies $\mathbb{E}[e^{|Z_1^{\bar{A},\bar{B}}|} \mid Z_0 = z] \leq \mathbb{E}[e^{|A^\kappa z + \mathcal{R}_*\operatorname{sat}_D(-\mathcal{R}_*^\dagger A^\kappa z)| + M_{\bar{V}} + M_{\bar{W}} + \|\mathcal{R}_*\|M_q D\epsilon} \mid Z_0 = z]$

holds, which is upper bounded by $\lambda(\epsilon)$ when $z \in \{z \mid |\mathcal{R}_*^\dagger A^\kappa z| > D\}$, and $\beta(\epsilon)$ when $z \in \{z \mid |\mathcal{R}_*^\dagger A^\kappa z| \leq D\}$. The conclusion in (8) follows via a modification of [18, Proposition 1] for systems exhibiting instability, with $V(\cdot) \leftarrow e^{(\cdot)}$.

Remark 6. The bound in (8) always provides upper bounds on the conditional expectation of $e^{|Z_\tau^{\bar{A},\bar{B}}|}$, and varies continuously as a function of the estimation error ϵ. However, if $\epsilon > 0$ is sufficiently small such that $\frac{-D}{\|\mathcal{R}_*^\dagger\|} + \|\mathcal{R}_*\| M_q D\epsilon + M_{\bar{V}} + M_{\bar{W}} < 0$ holds, then $\lambda(\epsilon) \in (0,1)$ will hold (under Assumption 4). In this scenario, the upper bound in (8) will asymptotically converge towards a constant, and can be qualitatively interpreted as a stability bound.

Using Lemma 2, we derive a probabilistic bound on the norm of the sub-sampled states $\bar{X}_\tau$ of the closed-loop system in (4) in Proposition 2. Note that the upper bound here depends on the function h, which is an arbitrarily chosen function over τ representing an upper bound on the estimation error for the sub-sampled parameter estimates $\bar{\theta}_\tau$ that holds after time step τ_0. The probability that this bound holds depends on the probability that the sub-sampled estimation error is bounded by h between τ and τ_0.

Proposition 2. *Suppose Assumptions 1, 2, and 3 hold, and let* $m_q, M_q > 0$ *satisfy Lemma 1. Suppose* $x_0 \in \mathbb{R}^n$. *Consider any function* $h : \{\tau_0, \tau_0+1, \ldots, \tau-1\} \to [0, m_q]$ *with* $\tau_0 \in \mathbb{N}$. *Fix* $\tau > \tau_0$ *and* $\gamma \in (0,1)$. *Then,*

$$|\bar{X}_\tau| < \ln\left(\frac{1}{\gamma}\right) + \ln\Bigg(\mathbb{E}[e^{|\bar{X}_{\tau_0}|}] \prod_{i=\tau_0}^{\tau-1} \lambda(h(i)) + \sum_{i=\tau_0}^{\tau-1} \beta(h(i)) \prod_{j=i+1}^{\tau-1} \lambda(h(j))\Bigg).$$

holds with at least probability $1 - \gamma - P(\bigcup_{i=\tau_0}^{\tau}\{\bar{\theta}_i \notin \bar{\mathcal{B}}_{h(i)}(\theta_*)\})$*, with* λ, β *defined in* (9) *and* (10) *respectively.*

We describe the key ideas for proving Proposition 2. Define the events $\mathcal{E}^2 = \{\bar{\theta}_{\tau_0} \in \bar{\mathcal{B}}_{h(\tau_0)}(\theta_*), \ldots, \bar{\theta}_{\tau-1} \in \bar{\mathcal{B}}_{h(\tau-1)}(\theta_*)\}$, and $\mathcal{E}_\gamma^1 = \{\mathbf{1}_{\mathcal{E}^2} e^{|\bar{X}_\tau|} < \frac{1}{\gamma}(\mathbb{E}[e^{|\bar{X}_{\tau_0}|}]\prod_{i=\tau_0}^{\tau-1}\lambda(h(i)) + \sum_{i=\tau_0}^{\tau-1}\beta(h(i))\prod_{j=i+1}^{\tau-1}\lambda(h(j)))\}$ for $\gamma \in (0,1)$. By employing an induction argument and making use of Lemma 2, we establish $\mathbb{E}[\mathbf{1}_{\mathcal{E}^2} e^{|\bar{X}_\tau|}] \leq \mathbb{E}[e^{|\bar{X}_{\tau_0}|}]\prod_{i=\tau_0}^{\tau-1}\lambda(h(i)) + \sum_{i=\tau_0}^{\tau-1}\beta(h(i))\prod_{j=i+1}^{\tau-1}\lambda(h(j))$. This result and Markov's inequality imply $P(\mathcal{E}_\gamma^1) \geq 1-\gamma$. The conclusion follows from $P\big(|\bar{X}_\tau| < \ln(\frac{1}{\gamma}) + \ln\big(\big[\mathbb{E}[e^{|\bar{X}_{\tau_0}|}]\prod_{i=\tau_0}^{\tau-1}\lambda(h(i)) + \sum_{i=\tau_0}^{\tau-1}\beta(h(i))\prod_{j=i+1}^{\tau-1}\lambda(h(j))\big]\big)\big) \geq P(\mathcal{E}_\gamma^1 \cap \mathcal{E}^2)$, and lower bounding $P(\mathcal{E}_\gamma^1 \cap \mathcal{E}^2)$ using a union bounding argument.

Combining Propositions 1 and 2, we are now ready to provide our main stability bound result in Theorem 2. It says, that given a failure probability δ and an estimation error parameter ϵ, we have a corresponding high probability upper bound on the sub-sampled states if ϵ is sufficiently small. Note from Remark 6 that when $\epsilon \in (0, \min(m_q, (\|\mathcal{R}_*\| M_q D)^{-1}(\frac{D}{\|\mathcal{R}_*^\dagger\|} - M_{\bar{V}} - M_{\bar{W}})))$, $\lambda(\epsilon) \in (0,1)$ holds, and therefore the right hand side of (11) is upper bounded by $\ln(2/\delta) + K(\epsilon, \delta/2, x_0) + \ln(\frac{\beta(\epsilon)}{1-\lambda(\epsilon)})$ over all τ, and asymptotically converges to $\ln(2/\delta) + \ln(\frac{\beta(\epsilon)}{1-\lambda(\epsilon)})$ as $\tau \to \infty$. Thus, it can be viewed as providing a parameterized (in ϵ) family of high probability stability bounds.

Theorem 2. *Suppose Assumptions 1, 2, 3, 4, and 5 hold, and let* $m_q, M_q > 0$ *satisfy Lemma 1. Suppose* $x_0 \in \mathbb{R}^n$. *Fix* $\delta \in (0,1)$ *and* $\epsilon \in (0, \min(m_q, (\|\mathcal{R}_*\| M_q D)^{-1}(\frac{D}{\|\mathcal{R}_*^\dagger\|} - M_{\bar{V}} - M_{\bar{W}})))$. *Then,*

$$|\bar{X}_\tau| < \left(\frac{2}{\delta}\right) + \ln\left(K(\epsilon, \delta/2, x_0)\lambda^\tau(\epsilon) + \frac{\beta(\epsilon)}{1-\lambda(\epsilon)}\right) \quad (11)$$

holds with probability at least $1-\delta$ *for all* $\tau \geq 0$*, where*

$$K(\epsilon, \delta, x_0) := e^{|x_0|}(e^{\|\mathcal{R}_*\| D + M_{\bar{V}} + M_{\bar{W}}}\lambda^{-1})^{\tau_0'(\epsilon,\delta,x_0)},$$
$$\tau_0'(\epsilon, \delta, x_0) := \min\{\tau \in \mathbb{N} \mid \kappa\tau \geq T_0(\delta, x_0), \; e(\kappa i, \delta, x_0) \leq \epsilon \text{ for all } i \geq \tau\},$$

with T_0, e, λ, β *defined in* (7), (6), (9), (10) *respectively.*

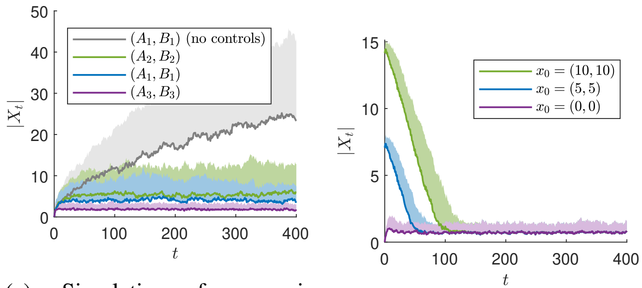
We describe the key ideas for proving Theorem 2. Firstly, from the dynamics (4), Assumption 2, and $\mathrm{sat}_D(\cdot) \leq D$, we have for all $\tau \geq 0$, $\mathbb{E}[e^{|\bar{X}_\tau|}] \leq e^{|x_0|}(e^{\|\mathcal{R}_*\| D + M_{\bar{V}} + M_{\bar{W}}})^\tau$ holds, which after applying Markov's inequality implies $|\bar{X}_\tau| < \ln(\frac{1}{\gamma}e^{|x_0|}(e^{\|\mathcal{R}_*\| D + M_{\bar{V}} + M_{\bar{W}}})^\tau)$ with probability at least $1-\gamma$ for all $\gamma \in (0,1)$. For the case where $\tau \leq \tau_0'(\epsilon, \delta, x_0)$, it can subsequently be deduced that $|\bar{X}_\tau| < \ln(2/\delta) + \ln(K(\epsilon, \delta/2)\lambda^\tau(\epsilon) + \frac{\beta(\epsilon)}{1-\lambda(\epsilon)})$ holds with probability $1-\delta/2$ after setting $\gamma \leftarrow \delta/2$ and some simplification, which implies that it holds with probability at least $1-\delta$. For the case where $\tau > \tau_0'(\epsilon, \delta, x_0)$, let $h(i) = e(\kappa i, \delta/2, x_0)$. Using Proposition 1, we have $P(\bigcup_{i=\tau_0'(\epsilon,\delta/2,x_0)}^{\tau}\{\bar{\theta}_i \notin \bar{\mathcal{B}}_{h(i)}(\theta_*)\}) \leq \delta/2$. After combining this with Proposition 2 where we set $\tau_0 \leftarrow \tau_0'(\epsilon, \delta, x_0)$, and then bounding $\mathbb{E}[e^{|\bar{X}_{\tau_0}|}]$ using our earlier mentioned anytime bound on $\mathbb{E}[e^{|\bar{X}_\tau|}]$, we find that $|\bar{X}_\tau| < \ln(2/\delta) + \ln(K(\epsilon, \delta/2)\lambda^\tau(\epsilon) + \frac{\beta(\epsilon)}{1-\lambda(\epsilon)})$ holds with probability $1-\delta$ after simplification. The conclusion follows by combining both cases.

Finally, we establish Theorem 1 by showing that under the Assumptions in Theorem 2, for all $\epsilon > 0$, there exist $L_1 > 0$ and a strictly increasing function $L_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\tau_0'(\epsilon, \delta) \leq L_2(|x_0|) + L_1 \ln(1/\delta)$ for all $x_0 \in \mathbb{R}^n$ and $\delta \in (0,1)$. This result is then applied to simplify Theorem 2 and yield Theorem 1. We defer the formal proof to the supplementary materials.

V. Numerical Examples

We demonstrated the effectiveness of our strategy by testing it on three different plants where $W_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \mathbf{I})$:

1) $A_1 = \begin{bmatrix} \cos(\pi/4) & \sin(\pi/4) \\ -\sin(\pi/4) & \cos(\pi/4) \end{bmatrix}$, $B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$;
2) $A_2 = \begin{bmatrix} \cos(-\pi/2) & \sin(-\pi/2) \\ -\sin(-\pi/2) & \cos(-\pi/2) \end{bmatrix}$, $B_2 = \begin{bmatrix} 0.3 \\ -0.5 \end{bmatrix}$;
3) $A_3 = \begin{bmatrix} 0.8\cos(\pi/4) & 0.8\sin(\pi/4) \\ -0.8\sin(\pi/4) & 0.8\cos(\pi/4) \end{bmatrix}$, $B_3 = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}$.



(a) Simulation for varying (A, B), and with no controls.

(b) Simulation for varying x_0.

Fig. 1: Median and 90th percentile of $|X_t|$ over 100 trials.

The algorithm parameters were also fixed to $U_{\max} \leftarrow 1$, $C \leftarrow 0.4$, $\kappa \leftarrow 2$, with $V_t \overset{\text{i.i.d.}}{\sim} \text{Uniform}(-C, C)$ and $(\bar{A}_0, \bar{B}_0)$ randomly selected (but fixed across all trials). We additionally simulated system (A_1, B_1) when it is uncontrolled (i.e. $U_t = 0$). The plots of the median and 90th percentiles of $|X_t|$ over 100 trials are shown in Fig. 1a. It can be seen that in all tests where Algorithm 1 is applied, at both the median and 90th percentile, $|X_t|$ seems to exhibit stable behaviour in the sense of boundedness, which is consistent with Theorem 1. This is in contrast to the case with no controls, where the median and 90th percentile plots have unbounded growth.

Secondly, we tested Algorithm 1 using the same algorithm parameters on (A_1, B_1) again, but with $\Sigma_W = 0.1\mathbf{I}$ and varying x_0. The plots are shown in Figure 1b. We see that regardless of the initial state, both the median and $90th$ percentile converge. In particular, if we focus individually on either the median or 90th percentile plots, and vary x_0, it appears that convergence occurs to the same steady state. Moreover, this convergence *seems* linear. This is consistent with the trends of the upper bound in Theorem 1.

VI. Conclusion

We proposed an excited CE control scheme for adaptive control of multi-dimensional, stochastic, linear systems subject to additive, i.i.d. unbounded stochastic disturbances, with positive upper bound constraints on the control magnitude. Moreover, we established a high probability stability bound on the κ-sub-sampled states of the closed-loop system. The stability of our strategy is verified in numerical examples.

This work can be extended in several directions. Assumption 2 can potentially be relaxed to linear systems where (A, B) are stabilizable, $\rho(A) \leq 1$, and eigenvalues of A on the unit circle have equal algebraic and geometric multiplicity, since the existence of mean square stabilizing controllers has been demonstrated [15]. However, controller design based on [15] involves a similarity transformation $T(A)$ that takes A to real Jordan form. Algorithm 1 could be modified to support such controllers, however, stability analysis would require inspection of the continuity properties of $T(A)$, and so we forgo the more general case in this work to simplify the exposition. Another interesting direction is output-feedback problems, since we only address the full-state feedback setting. Overcoming Assumption 4 is of interest, since in the non-adaptive setting, arbitrarily small controls are sufficient for stochastic stability [14]. Lastly, a more refined derivation of stability bounds than Theorems 2 and 1 is of interest, enabling a detailed analysis of the system behaviour beyond the trends observed here — for example, it may take into consideration the convergence of parameter estimates. We leave this to future work.

References

[1] T. Kargin, S. Lale, K. Azizzadenesheli, A. Anandkumar, and B. Hassibi, "Thompson sampling achieves $\tilde{O}(\sqrt{T})$ regret in linear quadratic control," in *Conf. Learn. Theory*, pp. 3235–3284, PMLR, 2022.

[2] M. Simchowitz and D. Foster, "Naive exploration is optimal for online lqr," in *Int. Conf. Mach. Learn.*, pp. 8937–8948, PMLR, 2020.

[3] Y. Li, S. Das, J. Shamma, and N. Li, "Safe adaptive learning-based control for constrained linear quadratic regulators with regret guarantees," *arXiv preprint arXiv:2111.00411*, 2021.

[4] M. S. Radenkovic and T. Altman, "Stochastic adaptive stabilization via extremum seeking in case of unknown control directions," *IEEE Trans. Autom. Control*, vol. 61, no. 11, pp. 3681–3686, 2016.

[5] M. S. Radenković and M. Krstić, "Extremum seeking-based perfect adaptive tracking of non-pe references despite nonvanishing variance of perturbation," *Automatica*, vol. 93, pp. 189–196, 2018.

[6] G. C. Goodwin, P. J. Ramadge, and P. E. Caines, "Discrete time stochastic adaptive control," *SIAM J. Control Optim.*, vol. 19, no. 6, pp. 829–853, 1981.

[7] L. Guo, "Self-convergence of weighted least-squares with applications to stochastic adaptive control," *IEEE Trans. Autom. Control*, vol. 41, no. 1, pp. 79–89, 1996.

[8] Y. Xiong, S. LeQuoc, and R. Cheng, "Stochastic adaptive control of multivariable systems with dead-zone nonlinearities," in *1993 American Control Conference*, pp. 489–493, IEEE, 1993.

[9] Z. Liu and C. Li, "Global stabilizability theorems on discrete-time nonlinear uncertain systems," *IEEE Transactions on Automatic Control*, 2022.

[10] F. Chaoui, F. Giri, and M. M'Saad, "Adaptive control of input-constrained type-1 plants stabilization and tracking," *Automatica*, vol. 37, no. 2, pp. 197–203, 2001.

[11] C. Zhang, "Adaptive control with input saturation constraints," in *Adaptive Control Nonsmooth Dyn. Syst.*, pp. 361–381, Springer, 2001.

[12] S. Siriya, J. Zhu, D. Nešić, and Y. Pu, "Learning-based adaptive control for stochastic linear systems with input constraints," *IEEE Control Systems Letters*, 2022.

[13] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Conference On Learning Theory*, pp. 439–473, PMLR, 2018.

[14] D. Chatterjee, F. Ramponi, P. Hokayem, and J. Lygeros, "On mean square boundedness of stochastic linear systems with bounded controls," *Systems & Control Letters*, vol. 61, no. 2, pp. 375–380, 2012.

[15] F. Ramponi, D. Chatterjee, A. Milias-Argeitis, P. Hokayem, and J. Lygeros, "Attaining mean square boundedness of a marginally stable stochastic linear system with a bounded control input," *IEEE Trans. Autom. Control*, vol. 55, no. 10, pp. 2414–2418, 2010.

[16] A. Tsiamis, I. Ziemann, N. Matni, and G. J. Pappas, "Statistical learning theory for control: A finite sample perspective," *arXiv preprint arXiv:2209.05423*, 2022.

[17] S. Siriya, J. Zhu, D. Nešić, and Y. Pu, "Stability bounds for learning-based adaptive control of discrete-time multi-dimensional stochastic linear systems with input constraints," *arXiv preprint arXiv:2304.00569*, 2023.

[18] D. Chatterjee and J. Lygeros, "On stability and performance of stochastic predictive control techniques," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 509–514, 2014.