Learning the Uncertainty Sets of Linear Control Systems via Set Membership: A Non-asymptotic Analysis

Yingying Li [*1] **Jing Yu** [*2] **Lauren Conger** [2] **Taylan Kargin** [2] **Adam Wierman** [2]

Abstract

This paper studies uncertainty set estimation for unknown linear systems. Uncertainty sets are crucial for the quality of robust control since they directly influence the conservativeness of the control design. Departing from the confidence region analysis of least squares estimation, this paper focuses on set membership estimation (SME). Though good numerical performances have attracted applications of SME in the control literature, the non-asymptotic convergence rate of SME for linear systems remains an open question. This paper provides the first convergence rate bounds for SME and discusses variations of SME under relaxed assumptions. We also provide numerical results demonstrating SME's practical promise.

1. Introduction

The problem of estimating unknown linear dynamical systems of the form $x_{t+1} = A^* x_t + B^* u_t + w_t$ with unknown parameters (A^*, B^*) has seen considerable progress recently (Sarker et al., 2023; Chen & Hazan, 2021; Simchowitz & Foster, 2020; Wagenmaker & Jamieson, 2020; Simchowitz et al., 2018; Dean et al., 2018; Abbasi-Yadkori & Szepesvári, 2011). Most literature focuses on the analysis of the least squares estimator (LSE) and its variants, where sharp bounds on the convergence rates for subGaussian disturbances w_t have been obtained (Simchowitz & Foster, 2020; Simchowitz et al., 2018). Building on this, there is a rapidly growing body of literature on "learning to control" unknown linear systems that leverages LSE to achieve various control objectives, such as stability and regret (Chang & Shahrampour, 2024; Lale et al., 2022; Simchowitz & Foster, 2020; Kargin et al., 2022; Mania et al., 2019; Dean et al., 2019b).

[*]Equal contribution [1]University of Illinois Urbana-Champaign [2]California Institute of Technology . Correspondence to: Yingying Li <yl101@illinois.edu>.

Proceedings of the 41^{st} International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

However, for successful application of learning-based control methods to safety-critical applications, it is crucial to quantify the uncertainties of the estimated system and to robustly satisfy safety constraints and stability despite these uncertainties (Wabersich et al., 2023; Brunke et al., 2022). A promising framework for achieving this is to estimate the uncertainty set of the unknown system parameters and to utilize robust controllers to satisfy the robust constraints under any parameters in the uncertainty set (Brunke et al., 2022; Hewing et al., 2020). Uncertainty set estimation is crucial for the success of robust control: on the one hand, too large of an uncertainty set gives rise to over-conservative control actions, resulting in degraded performance; on the other hand, if the uncertainty set is underestimated and fails to contain the true system, the resulting controller may lead to unsafe behaviors (Brunke et al., 2022; Petrik & Russel, 2019).

To estimate uncertainty sets, a popular method is to construct LSE's confidence regions (Dean et al., 2019b; Simchowitz & Foster, 2020). However, this approach yields a confidence region for a point estimate rather than directly estimating the uncertainty set of the model. Further, the confidence regions are usually derived from concentration inequalities, which allows convergence rate analysis but may suffer conservative constant factors (Petrik & Russel, 2019; Simchowitz & Foster, 2020).

In this paper, we instead focus on a direct uncertainty set estimation method: set membership estimation (SME), which estimates the uncertainty set without relying on the concentration inequalities underlying the approaches based on LSE. SME has a long history in the control community (Yu et al., 2023b; Lauricella & Fagiano, 2020; Lorenzen et al., 2019; Livstone & Dahleh, 1996; Fogel & Huang, 1982; Bertsekas, 1971). SME has primarily been proposed for scenarios with bounded disturbances, which is common in safety-critical systems, e.g. power systems (Qi et al., 2012), unmanned aerial vehicles (UAV) (Benevides et al., 2022; Narendra & Annaswamy, 1986), and building control (Zhang et al., 2016). Further, the bounded disturbance is a standard assumption in constrained control, such as robust (adaptive) constrained control (Lu & Cannon, 2023; Lorenzen et al., 2019; Dean et al., 2019b) and online constrained control (Li et al., 2021a; Liu et al., 2023).

Consequently, SME has been widely adopted in the robust (adaptive) constrained control literature (Lorenzen et al., 2019; Bujarbaruah et al., 2020; Zhang et al., 2021; Parsi et al., 2020b;a; Sasfi et al., 2022) and the online control literature (Ho et al., 2021; Yu et al., 2023b; Yeh et al., 2022; Yu et al., 2023a). Figure 1 provides a toy example illustrating SME's promising performance under bounded disturbances.

On the theory side, the convergence analysis of SME generally considers a simple regression problem: $y_t = \theta^* x_t + w_t$ with a deterministic sequence of x_t and bounded i.i.d. disturbances w_t (Bai et al., 1998; Akçay, 2004; Kitamura et al., 2005; Bai et al., 1995; Eising et al., 2022). This regression problem does not capture the correlation between x_t and the history $w_{t-1}, \ldots, w_0$ in the dynamical systems. This issue was largely overlooked in the vast literature of empirical algorithm design related to SME (for example, see (Lorenzen et al., 2019; Köhler et al., 2019), etc.). It is not until recently that (Lu et al., 2019) provide the first *asymptotic convergence* guarantees for SME in linear systems. However, the *non-asymptotic convergence rate* still remains open for SME in linear dynamical systems.

Contributions. This paper tackles the open question above by providing non-asymptotic bounds on the convergence rates of SME for linear systems. To the best of our knowledge, this is the first convergence rate analysis of SME for dynamical systems in the literature.

We consider two scenarios in our analysis. Firstly, when a *tight* bound $\mathcal{W}$ on the support of w_t is known, we provide an instance-dependent convergence rate for SME. Interestingly, for several common distributions of w_t, SME enjoys a convergence rate $\tilde{O}(n_x^{1.5}(n_x+n_u)^2/T)$, which is faster than the LSE's error bound $O(\frac{\sqrt{n_x+n_u}}{\sqrt{T}})$ in terms of the number of samples T but is worse in terms of the dependence on state and control dimensions n_x, n_u. The improved convergence rate of SME with respect to T is enabled by leveraging the additional boundedness property of w_t, which is a common assumption in robust constrained control but is not utilized in LSE's analysis. Secondly, when a tight bound of w_t is *unknown*, we introduce a UCB-SME algorithm that learns conservative upper bounds of w_t from data and constructs uncertainty sets based on the conservative upper bounds. We also provide a convergence rate of UCB-SME, which has the same dependence on T but has worse dependence on n_x by a factor of $\sqrt{n_x}$ compared with the convergence rate with a known tight bound.

Our estimation error bound relies on a novel construction of an event sequence based on designing a sequence of stopping times. This construction, together with the BMSB condition in (Simchowitz et al., 2018), addresses the challenge caused by the correlation between x_t, u_t, and the history disturbances (see the proof of Theorem 3.1 for more details).

Moreover, our results lay a foundation for future non-asymptotic analysis of control designs based on SME. To illustrate this, we apply our results to robust-adaptive model predictive control and robust system-level-synthesis (SLS) and discuss the novel non-asymptotic guarantees enabled by our convergence rates of SME.

Finally, we conduct extensive simulations to compare the numerical behaviors of SME, UCB-SME, and LSE's confidence regions, which demonstrates the promising performance of SME and UCB-SME.

2. Problem Formulation and Preliminaries

2.1. Problem Formulation

This paper focuses on the identification of uncertainty sets of unknown system parameters in the linear dynamical system:

$$x_{t+1} = A^* x_t + B^* u_t + w_t \tag{1}$$

where A^*, B^* are the unknown system parameters, $x_t \in \mathbb{R}^{n_x}, u_t \in \mathbb{R}^{n_u}$. For notational simplicity, we define $\theta^* = (A^*, B^*)$ by matrix concatenation and $z_t = (x_t^\top, u_t^\top)^\top \in \mathbb{R}^{n_z}$ by vector concatenation, where $n_z = n_x + n_u$. Accordingly, the system (1) can be written as $x_{t+1} = \theta^* z_t + w_t$.

The goal of the uncertainty set identification problem is to determine a set Θ_T that contains the true parameters $\theta^* = (A^*, B^*)$ based on a sequence of data $\{x_t, u_t, x_{t+1}\}_{t=0}^{T-1}$. Set Θ_T is called an uncertainty set since it captures the remaining uncertainty on the system model after the revelation of the data sequence $\{x_t, u_t, x_{t+1}\}_{t=0}^{T-1}$.

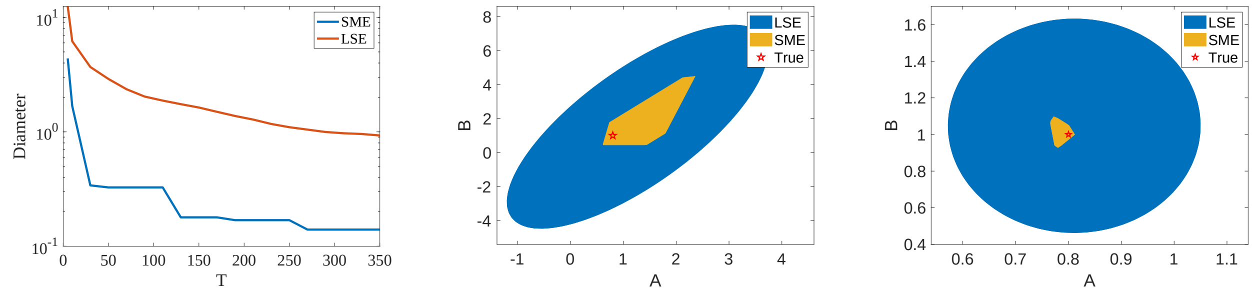
Uncertainty sets play an important role in robust control, where one aims to achieve robust constraint satisfaction (Lorenzen et al., 2019; Lu & Cannon, 2023), robust objective optimization (Wu et al., 2013), and/or robust stability (Tu, 2019) for any model in the uncertainty set.[1] Therefore, the diameter of the uncertainty sets heavily influences the conservativeness of robust controllers and thus the control performance. Formally, we define the diameter as follows.

Definition 2.1 (Diameter of a set of matrices)**.** Consider a set $\mathbb{S}$ of matrices $\theta \in \mathbb{R}^{n_x \times n_z}$. We define the diameter of $\mathbb{S}$ in Frobenius norm as $\text{diam}(\mathbb{S}) = \sup_{\theta,\theta' \in \mathbb{S}} \|\theta - \theta'\|_F$.

2.2. Set Membership Estimation (SME)

In this section, we review set membership estimation (SME), which is an uncertainty set identification method that has been studied in the control literature for decades (Lu &

[1] In addition to model uncertainties, robust control may also consider other system uncertainties, e.g., disturbances, measurement noises, etc.



(a) Uncertainty sets' diameters (b) $T = 5$ (c) $T = 250$

Figure 1. A visualized toy example of uncertainty set comparison between SME in (2) and LSE confidence regions in (Simchowitz & Foster, 2020; Abbasi-Yadkori & Szepesvári, 2011) for a one-dimensional system $x_{t+1} = A^* x_t + B^* u_t + w_t$, with w_t, $u_t \in [-1, 1]$ generated i.i.d. from a truncated Gaussian distribution. Detailed experiment settings are in Appendix I. **Figure (a)** compares the diameters of the uncertainty sets from SME and LSE 90% confidence bounds. **Figure (b) and (c)** visualize the the uncertainty sets after $T = 5$ and $T = 250$ data points.

Cannon, 2023; Bertsekas, 1971). SME primarily focuses on systems with *bounded* disturbances, i.e. $w_t \in \mathcal{W}$ for some bounded $\mathcal{W}$ for all $t \geq 0$. When $\mathcal{W}$ is known, SME computes an uncertainty/membership set by

$$\Theta_T = \bigcap_{t=0}^{T-1} \{\hat{\theta} : x_{t+1} - \hat{\theta} z_t \in \mathcal{W}\}. \tag{2}$$

It is straightforward to see that $\theta^* \in \Theta_T$ when $w_t \in \mathcal{W}$.

The bounded disturbance assumption may seem restrictive, considering that the uncertainty set identification based on the confidence region of LSE only requires subGaussian disturbances (Simchowitz et al., 2018). However, in many control applications, it is reasonable and common to assume bounded w_t. For example, bounded disturbances is a standard assumption in the robust constrained control literature, such as robust constrained LQR (Lu & Cannon, 2023; Lorenzen et al., 2019; Lu et al., 2019; Dean et al., 2019b), and online constrained control of linear systems (Liu et al., 2023; Li et al., 2021a). This is different from unconstrained control, where unbounded subGaussian disturbances are usually considered (Tu, 2019). The difference in the disturbance formulation is largely motivated by the applications: constrained control is mostly applied to safety-critical applications, where the disturbances are usually bounded. For example, in UAV and flight control, the disturbances are mostly caused by wind gusts, and wind disturbances are bounded in practice (Benevides et al., 2022; Narendra & Annaswamy, 1986). Similarly, in building thermal control, the disturbances are caused by external heat exchanges, which are also bounded (Zhang et al., 2016).

Ideally, one hopes that Θ_T converges to the singleton of the true model $\{\theta^*\}$ or at least a small neighborhood of θ^*. This usually calls for additional assumptions, such as the persistent excitation property on the observed data and additional stochastic properties on w_t. In this paper, we consider the following assumptions to establish convergence rate bounds on the diameter of Θ_T, which, to the best of our knowledge, is the first non-asymptotic guarantee of SME for linear dynamical systems.

The first assumption formalizes the bounded disturbance assumption discussed above and introduces stochastic properties of w_t for analytical purposes.

Assumption 2.2 (Bounded i.i.d. disturbances)**.** *The disturbances are box-constrained, $w_t \in \mathcal{W} := \{w \in \mathbb{R}^{n_x} : \|w\|_\infty \leq w_{\max}\}$ for all $t \geq 0$. Further, w_t is i.i.d., has zero mean and positive definite covariance matrix Σ_w.*

Assumption 2.2 is common in SME literature, e.g. (Akçay, 2004; Lu et al., 2019; Eising et al., 2022). In terms of generality, boundedness is essential for SME. The stochastic properties, such as i.i.d., zero mean, positive definite covariance, are standard in the recent learning-based control literature and allow the use of statistical tools utilized and developed in the recent literature for non-asymptotic analysis. Besides, it is worth mentioning that SME still works in non-stochastic settings. In particular, as long as $w_t \in \mathcal{W}$, even without the stochastic properties in Assumption 2.2, the SME algorithm (2) still generates a valid uncertainty set that contains θ^*. It is an interesting future direction to study the convergence rate of SME without assuming stochastic disturbances.

Next, we introduce the assumptions on u_t, which relies on the block-martingale small-ball (BMSB) condition proposed in (Simchowitz et al., 2018). It can be shown that the BMSB guarantees persistent excitation (PE) with high probability under proper conditions (see Proposition 2.5 in (Simchowitz et al., 2018) and Lemma 4.1). The PE condition requires that z_t explores all directions, which is essential for system identification (Narendra & Annaswamy, 1987).

Definition 2.3 (Persistent excitation)**.** There exists $\alpha > 0$ and $m \in \mathbb{N}+$, such that for any $t_0 \geq 0$,

$$\frac{1}{m}\sum_{t=t_0}^{t_0+m-1}\begin{pmatrix}x_t\\u_t\end{pmatrix}(x_t^\top, u_t^\top) \succeq \alpha^2 I_{n_x+n_u}.$$

Definition 2.4 (BMSB (Simchowitz et al., 2018))**.** Consider a filtration $\{\mathcal{F}_t\}_{t\geq 1}$ and an $\{\mathcal{F}_t\}_{t\geq 1}$-adapted random process $\{Z_t\}_{t\geq 1}$ in $\mathbb{R}^d$. $\{Z_t\}_{t\geq 1}$ satisfies the (k, Γ_{sb}, p)-block martingale small-ball (BMSB) condition for $k > 0$, a positive definite Γ_{sb}, and $0 \leq p \leq 1$, if the following holds: for any fixed $\lambda \in \mathbb{R}^d$ with $\|\lambda\|_2 = 1$, we have $\frac{1}{k}\sum_{i=1}^k \mathbb{P}(|\lambda^\top Z_{t+i}| \geq \sqrt{\lambda^\top \Gamma_{sb}\lambda} \mid \mathcal{F}_t) \geq p$ for all $t \geq 1$.

The following is the assumption on u_t.

Assumption 2.5 (BMSB and boundedness)**.** *With filtration $\mathcal{F}_t = \mathcal{F}(w_0, \dots, w_{t-1}, z_0, \dots, z_t)$, the $\mathcal{F}_t$-adapted stochastic process $\{z_t\}_{t\geq 0}$ satisfies $(1, \sigma_z^2 I_{n_z}, p_z)$-BMSB for some $\sigma_z, p_z > 0$. Besides, there exists $b_z \geq 0$ such that $\|z_t\|_2 \leq b_z$ almost surely for all $t \geq 0$.*

Assumption 2.5 requires u_t to guarantee both BMSB and bounded z_t. This can be satisfied by several robust (adaptive) constrained control policies, such as robust (adaptive) model predictive control (MPC) (Lu & Cannon, 2023; Lorenzen et al., 2019; Lu et al., 2019), system level synthesis (SLS) (Dean et al., 2019b), and control barrier functions (CBF) (Lopez et al., 2020). In the following, we briefly discuss robust (adaptive) MPC as an example. SLS and CBF can be similarly shown to satisfy Assumption 2.5.

Example 1 (Robust (adaptive) MPC)**.** Robust MPC is a popular method for the robust constrained control (Rawlings & Mayne, 2009), which aims to optimize the control objective while satisfying robust safety constraints,

$$z_t \in \mathbb{Z}_{\text{safe}}, \text{ where } x_{t+1} = \theta z_t + w_t, \forall\, \theta \in \Theta_0, w_t \in \mathcal{W}, \quad (3)$$

where Θ_0 is an initial uncertainty set known a priori, and the safety constraint $\mathbb{Z}_{\text{safe}}$ is usually bounded. The robust MPC policy, denoted by $u_t = \pi_{\text{RMPC}}(x_t; \Theta_0, \mathcal{W})$, satisfies the constraints (3) for any $\theta \in \Theta_0$. Therefore, it naturally guarantees bounded z_t under the true θ^*. Further, as shown in (Li et al., 2023), BMSB can be achieved by adding a random disturbance, i.e. $u_t = \pi_{\text{RMPC}}(x_t; \Theta_0, \mathcal{W}) + \eta_t$, where η_t is i.i.d., bounded, and has positive definite covariance. Therefore, the randomly perturbed robust MPC can satisfy Assumption 2.5.[2] Robust adaptive MPC is based on the same control design, $u_t = \pi_{\text{RMPC}}(x_t; \Theta_t, \mathcal{W})$, but utilizes adaptively updated uncertainty sets Θ_t. Notice that Θ_t is usually updated by SME in the literature of robust adaptive MPC (Lorenzen et al., 2019; Lu & Cannon, 2023; Köhler et al., 2019).

[2]Strictly speaking, robust MPC has to be more conservative to satisfy (3) under the additional noise η_t (Li et al., 2023).

We also note that BMSB and bounded z_t with high probability are assumed in LSE literature (Theorem 2.4 (Simchowitz et al., 2018)), and bounded z_t with high probability under subGaussian disturbances corresponds to bounded z_t under bounded disturbances for linear systems (see bounded-input-bounded-output stability in Sec. 9 of (Hespanha, 2018)).

Finally, we assume that the bound $w_{\max}$ on w_t is tight in all directions, which is common in the literature on SME analysis (Bai et al., 1998; Akçay, 2004; Lu et al., 2019).

Assumption 2.6 (Tight bound on w_t)**.** *For any $\epsilon > 0$, there exists $q_w(\epsilon) > 0$, such that for any $1 \leq j \leq n$, we have*

$$\min(\mathbb{P}(w_t^j \leq \epsilon - w_{\max}), \mathbb{P}(w_t^j \geq w_{\max} - \epsilon)) \geq q_w(\epsilon),$$

where w_t^j denotes the jth entry of vector w_t. Without loss of generality, we can further assume $q_w(\epsilon)$ to be non-decreasing with ϵ and $q_w(2w_{\max}) = 1$.[3]

In essence, Assumption 2.6 requires that a *hyper-cubic* $\mathcal{W} = \{w : \|w\|_\infty \leq w_{\max}\}$ should be *tight* on the support of w_t in all coordinate directions, that is, there exists a positive probability $q_w(\epsilon)$ such that w_t visits an ϵ-neighborhood of $w_{\max}$ and $-w_{\max}$, respectively, on all coordinates.

When the support of w_t is indeed $\mathcal{W} = \{w : \|w\|_\infty \leq w_{\max}\}$, many common distributions enjoy $q_w(\epsilon) \geq \Omega(\epsilon)$.[4] For example, for the uniform distribution on $\mathcal{W}$, we have $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}}$; for the truncated Gaussian distribution with zero mean, $\sigma_w^2 I_n$ covariance, and truncated region $\mathcal{W}$, we have $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}\sigma_w}\exp(\frac{-w_{\max}^2}{2\sigma_w^2})$; and for the uniform distribution on the boundary of $\mathcal{W}$ (a generalization of Rademacher distribution), we have $q_w(\epsilon) \geq \frac{1}{2n_x} \geq \Omega(\epsilon)$ (see Appendix C.2 for more details).

However, knowing a tight bound on the support of w_t can be challenging in practice. Therefore, we will discuss how to relax this assumption and learn a tight bound from data in Section 3.2.

Further, the requirement of a hyper-cubic $\mathcal{W}$ can be restrictive because different entries of disturbances may have different magnitudes, resulting in a hyper-rectangular support that violates Assumption 2.6. Our follow-up work relaxes this assumption and generalizes the results in this paper.

3. Set Membership Convergence Analysis

3.1. Convergence Rate of SME with Known $w_{\max}$

We now present the main result (Theorem 3.1) of this paper, which is a non-asymptotic bound on the estimation error of SME given bounded i.i.d. stochastic disturbances.

[3]This is because $\mathbb{P}(w_t^j \leq \epsilon - w_{\max})$ and $\mathbb{P}(w_t^j \geq w_{\max} - \epsilon)$ are non-decreasing with ϵ, and $\mathbb{P}(w_t^j \geq -w_{\max}) = \mathbb{P}(w_t^j \leq w_{\max}) = 1$ by Assumption 2.2.

[4]The $\Omega(\cdot)$ notation is the lower bound version of $O(\cdot)$.

Theorem 3.1 (Convergence rate of SME)**.** *For any $m > 0$ any $\delta > 0$, when $T > m$, we have*

$$\mathbb{P}(\text{diam}(\Theta_T) > \delta) \leq \underbrace{\frac{T}{m}\tilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m)}_{\mathbb{T}_1} + \underbrace{\tilde{O}((n_x n_z)^{2.5})a_4^{n_x n_z}\left(1 - q_w\left(\frac{a_1\delta}{4\sqrt{n_x}}\right)\right)^{\lceil T/m \rceil}}_{\mathbb{T}_2(\delta)}$$

where $a_1 = \frac{\sigma_z p_z}{4}$, $a_2 = \frac{64b_z^2}{\sigma_z^2 p_z^2}$, $a_3 = \frac{p_z^2}{8}$, $a_4 = \frac{4b_z\sqrt{n_x}}{a_1}$, p_z, σ_z, b_z are defined in Assumption 2.5, $\lceil\cdot\rceil$ denotes the ceiling function, and $\text{diam}(\cdot)$ *is defined in Definition 2.1, the factors hidden in $\tilde{O}(\cdot)$ are provided in Appendix D.4.*

Theorem 3.1 provides an upper bound on the "failure" probability of SME, i.e., the probability that the diameter of the uncertainty set is larger than δ. In this bound, $\mathbb{T}_1$ decays exponentially with m, so for any small $\epsilon > 0$, m can be chosen such that $\mathbb{T}_1 \leq \epsilon$, which indicates $m \geq O(n_z + \log T + \log(1/\epsilon))$. For any $\delta > 0$, $\mathbb{T}_2(\delta)$ decays exponentially with the number of data points T and involves a distribution-dependent function $q_w(\cdot)$, which characterizes how likely it is for w_t to visit the boundary of $\mathcal{W}$ as defined in Assumption 2.6. To ensure the probability upper bound in Theorem 3.1 to be less than 1, one can choose $m = O(\log T)$ and a large enough T such that $T \geq O(m) = O(\log(T))$. If w_t is more likely to visit the boundary, (a larger $q_w(\cdot)$), then SME is less likely to generate an uncertainty set with a diameter bigger than δ.

Estimation error bounds when $q_w(\epsilon) = \Omega(\epsilon)$. To provide intuition for $\mathbb{T}_2(\delta)$ and discuss the estimation error bound in Theorem 3.1 more explicitly, we consider distributions satisfying $q_w(\epsilon) = \Omega(\epsilon)$ for all $\epsilon > 0$. Notice that several common distributions satisfy this additional requirement, such as uniform distribution and truncated Gaussian distribution as discussed after Assumption 2.6.

Corollary 3.2 (Estimation error bound when $q_w(\epsilon) = \Omega(\epsilon)$)**.** *For any $\epsilon > 0$, let*

$$m \geq O(n_z + \log T + \log(1/\epsilon))$$

in the following.[5] *If w_t is generated i.i.d. by a distribution satisfying $q_w(\epsilon) = \Omega(\epsilon)$ for all $\epsilon > 0$, then with probability at least $1 - 2\epsilon$, for any $\hat{\theta}_T \in \Theta_T$, we have*

$$\|\hat{\theta}_T - \theta^*\|_F \leq \text{diam}(\Theta_T) \leq \tilde{O}\left(\frac{n_x^{1.5}(n_x + n_u)^2}{T}\right).$$

Corollary 3.2 indicates that the estimation error of any point in the uncertainty set Θ_T can be bounded by $\tilde{O}\left(\frac{n_x^{1.5}(n_x+n_u)^2}{T}\right)$ when $q_w(\epsilon) \geq \Omega(\epsilon)$.

[5] A detailed formula is provided in Appendix E.

Dynamical systems without control inputs. SME also applies to dynamical systems with no control inputs, i.e., $x_{t+1} = A^* x_t + w_t$, where the uncertainty set of A^* can be computed by $\mathbb{A}_T = \bigcap_{t=0}^{T-1}\{\hat{A} : \|x_{t+1} - \hat{A}x_t\|_\infty \leq w_{\max}\}$. Its convergence rate can be similarly derived via the proof of Theorem 3.1.

Corollary 3.3 (Convergence rate with $B^* = 0$ (informal))**.** *For stable A^*, for any $m > 0, \delta > 0, T > m$, we have*

$$\mathbb{P}(\text{diam}(\mathbb{A}_T) > \delta) \leq \frac{T}{m}\tilde{O}(n_x^{2.5})a_2^{n_x}\exp(-a_3 m) + \tilde{O}(n_x^5)a_4^{n_x^2}(1 - q_w(\frac{a_1\delta}{4\sqrt{n}_x}))^{\lceil T/m \rceil}$$

Consequently, when $q_w(\epsilon) = \Omega(\epsilon)$, e.g. uniform or truncated Gaussian, we have $\text{diam}(\mathbb{A}_T) \leq \tilde{O}(n_x^{3.5}/T)$.

Note that (Simchowitz et al., 2018) have shown a lower bound $\Omega(\sqrt{n}_x/\sqrt{T})$ for the estimation of linear systems with no control inputs when w_t follows an (unbounded) Gaussian distribution. Interestingly, Corollary 3.3 reveals that, for some bounded-support distributions of w_t, e.g. Uniform and truncated Gaussian, SME is able to converge at a faster rate $\tilde{O}(1/T)$ in terms of the sample size T. This does not conflict with the lower bound in (Simchowitz et al., 2018) because SME's rate only holds for bounded disturbances. In fact, from (2), it is straightforward to see that SME does not even converge under Gaussian disturbances. Therefore, SME is mostly useful in applications with bounded disturbances, e.g. robust constrained control, safety-critical systems, etc., while LSE's confidence regions are preferred for unbounded disturbances.

Lastly, Corollary 3.3 shows that SME's convergence rate has a poor dependence with respect to n_x: $\tilde{O}(n_x^{3.5})$. This is likely a proof artifact because we do not observe such poor dimension scaling in simulation (see Figure 3). It is left as future work to refine the dimension dependence.

3.2. SME with Unknown $w_{\max}$

Next, we discuss the convergence rates of SME without knowing a tight bound $w_{\max}$ in three steps: 1) only knowing a conservative upper bound of $w_{\max}$, 2) learning $w_{\max}$ from data, and 3) a variant of SME that converges without prior knowledge of w_{max}.

1) SME with a conservative upper bound for $w_{\max}$. In many practical scenarios, it is easier to obtain an over-estimation of the range of the disturbances instead of a tight upper bound, i.e., $\hat{w}_{\max} \geq w_{\max}$. In this case, we can show that the uncertainty set converges to a small neighborhood around θ^* of size $O(\sqrt{n_x}(\hat{w}_{\max} - w_{\max}))$ at the same convergence rate as Theorem 3.1.

Theorem 3.4 (Convservative bound on $w_{\max}$)**.** *When $w_{\max}$ in Assumption 2.6 is unknown but an upper bound $\hat{w}_{\max} \geq$*

$w_{\max}$ is known, consider the following SME algorithm:

$$\hat{\Theta}_T(\hat{w}_{\max}) = \bigcap_{t=0}^{T-1} \{\hat{\theta} : \|x_{t+1} - \hat{\theta} z_t\|_\infty \leq \hat{w}_{\max}\},$$

For any $m > 0$, $\delta > 0$, $T > m$, we have

$$\mathbb{P}(\text{diam}(\hat{\Theta}_T) > \delta + a_5\sqrt{n_x}(\hat{w}_{\max} - w_{\max})) \leq \mathbb{T}_1 + \mathbb{T}_2(\delta)$$

where $a_5 = \frac{4}{a_1}$, $\mathbb{T}_1, \mathbb{T}_2(\delta)$ are defined in Theorem 3.1.

2) Learning $w_{\max}$. When $w_{\max}$ is not accurately known, we can try to learn it from the data. Let's first consider the learning algorithm studied in (Bai et al., 1998).

$$\bar{w}_{\max}^{(T)} = \min_\theta \max_{0 \leq t \leq T-1} \|x_{t+1} - \theta z_t\|_\infty. \tag{4}$$

Though algorithm (4) cannot provide an upper bound on $w_{\max}$ under finite samples because $\bar{w}_{\max}^{(T)} \leq w_{\max}$ for finite T,[6] it can be shown that $\bar{w}_{\max}^{(T)}$ converges to $w_{\max}$ as $T \to +\infty$. The convergence for linear regression has been established in (Bai et al., 1998). The following theorem establishes the convergence and convergence rate of algorithm (4) for linear dynamical systems. Based on this convergence rate, we will design an online learning algorithm (5) that generates converging upper bounds of $w_{\max}$.

Theorem 3.5. *The estimation $\bar{w}_{\max}^{(T)}$ of $w_{\max}$ satisfies:*

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \underbrace{b_z \text{diam}(\Theta_T)}_{\mathbb{T}_3} + \underbrace{w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty}_{\mathbb{T}_4}$$

Therefore, for any $\delta > 0$,

$$\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) \leq \mathbb{T}_1 + \mathbb{T}_2\left(\frac{\delta}{2b_z}\right) + \mathbb{T}_5\left(\frac{\delta}{2}\right),$$

where $\mathbb{T}_5(\delta) = (1 - q_w(\delta))^T$.

Notice that $\mathbb{T}_4$ is the smallest possible learning error of $w_{\max}$ from history w_t, which can be achieved if one can directly measure w_t. However, with unknown θ^*, it is challenging to measure/compute w_t exactly, then Theorem 3.5 shows that the learning error of $w_{\max}$ has an additional term $\mathbb{T}_3$ that depends on the uncertainty around θ^*. Therefore, the convergence rate of $\bar{w}_{\max}^{(T)}$ can be obtained by our non-asymptotic analysis of SME in Theorem 3.1.

Further, when $q_w(\epsilon) = \Omega(\epsilon)$, the convergence rate of $\bar{w}_{\max}^{(T)}$ can be explicitly bounded by $\tilde{O}(n_x^{1.5} n_z^2/T)$, which is of the same order as the convergence rate of the diameter of Θ_T.

Corollary 3.6. *For any $0 < \epsilon < 1/3$ and any $T \geq 1$, there exists $\delta_T > 0$ satisfying $\lim_{T\to\infty} \delta_T = 0$ such that*

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \delta_T$$

with probability at least $1 - 3\epsilon$.

In particular, when $q_w(\delta) = O(\delta)$, with probability $1 - 3\epsilon$,

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq \delta_T = \tilde{O}(n_x^{1.5} n_z^2/T)$$

3) SME with unknown $w_{\max}$. Unfortunately, $\bar{w}_{\max}^{(T)}$ cannot be directly applied to SME because $\bar{w}_{\max}^{(T)} \leq w_{\max}$, which may cause $\theta^* \notin \hat{\Theta}_T(\bar{w}_{\max}^{(T)})$. However, by leveraging our convergence rate bound in Theorem 3.5, we can construct an upper confidence bound (UCB) of $w_{\max}$ and a corresponding UCB-SME algorithm:

$$\hat{w}_{\max}^{(T)} = \bar{w}_{\max}^{(T)} + \delta_T, \quad \hat{\Theta}_T^{\text{ucb}} = \hat{\Theta}_T(\hat{w}_{\max}^{(T)}), \tag{5}$$

where δ_T is defined in Corollary 3.6.

Then, by combining Theorem 3.4 and Corollary 3.6, we can verify the well-definedness of UCB-SME and obtain its convergence rate.

Theorem 3.7. *For any $0 < \epsilon < 1/3$, any $T \geq 1$, with probability at least $1 - 3\epsilon$, we have*

$$\theta^* \in \hat{\Theta}_T^{\text{ucb}}, \quad \text{diam}(\hat{\Theta}_T^{\text{ucb}}) \leq O(\sqrt{n_x}\delta_T).$$

In particular, if $q_w(\epsilon) = \Omega(\epsilon)$, then $\text{diam}(\hat{\Theta}_T^{\text{ucb}}) \leq O(n_x^2 n_z^2/T)$ with probability at least $1 - 3\epsilon$.

Notice that UCB-SME converges at the same rate in terms of T but $\sqrt{n_x}$-worse in terms of dimensionality when compared with SME knowing a tight bound $w_{\max}$.

Remark 3.8 (Computation complexity)**.** *SME can be computed by linear programming since all constraints are linear in (2). Further, UCB-SME can also be computed by linear programming because (4) can be reformulated as a linear program. However, the number of constraints for SME and UCB-SME increases linearly with T. To address the computation issue of SME, many computationally efficient algorithms have been proposed based on approximations of (2), e.g. (Lu et al., 2019; Yeh et al., 2022; Bai et al., 1995). The convergence rates of these approximate algorithms are unknown and how to design computationally efficient UCB-SME remains open.*

4. Proof Sketch of Theorem 3.1

The major technical novelty of this paper is the proof of Theorem 3.1, thus we describe the key ideas here. The complete proof is provided in Appendix D. For ease of notation and without loss of generality, we assume T/m is an integer in the following.

Specifically, we first define a set Γ_T on the model estimation error $\gamma = \hat{\theta} - \theta^*$ by leveraging the observation that $x_{s+1} - \hat{\theta} z_s = w_s - (\hat{\theta} - \theta^*) z_s$,

$$\Gamma_t = \bigcap_{s=0}^{t-1} \{\gamma : \|w_s - \gamma z_s\|_\infty \leq w_{\max}\}, \quad \forall\, t \geq 0. \tag{6}$$

[6] If SME does not use an upper bound on $w_{\max}$, the generated uncertainty set may not contain the true parameter θ^*.

Notice that $\Theta_t = \theta_* + \Gamma_t$, so $\text{diam}(\Theta_t) = \text{diam}(\Gamma_t)$, and

$$\text{diam}(\Gamma_t) = \sup_{\gamma,\gamma' \in \Gamma_t} \|\gamma - \gamma'\|_F \leq 2 \sup_{\gamma \in \Gamma_t} \|\gamma\|_F.$$

Thus, we can define $\mathcal{E}_1 \coloneqq \{\exists\, \gamma \in \Gamma_T, \text{ s.t. } \|\gamma\|_F \geq \frac{\delta}{2}\}$ such that $\mathbb{P}(\text{diam}(\Theta_T) > \delta) \leq \mathbb{P}(\mathcal{E}_1)$.

Next, we define an event $\mathcal{E}_2$ below, which is essentially PE on every time segments $km + 1 \leq t \leq km + m$ for $k \geq 0$, where the choice of m will be specified later.

$$\mathcal{E}_2 = \left\{ \frac{1}{m} \sum_{s=1}^{m} z_{km+s} z_{km+s}^\top \succeq a_1^2 I_{n_z}, \forall\, 0 \leq k \leq \left\lceil \frac{T}{m} \right\rceil - 1 \right\}$$

where $a_1 = \frac{\sigma_z p_z}{4}$.

Now, by dividing the event $\mathcal{E}_1$ based on $\mathcal{E}_2$, we obtain

$$\mathbb{P}(\text{diam}(\Theta_T) > \delta) \leq \mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2^c) + \mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2).$$

The proof can be completed by establishing the following bounds on $\mathbb{P}(\mathcal{E}_2^c)$ and $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2)$.

Lemma 4.1 (Bound on $\mathbb{P}(\mathcal{E}_2^c)$)**.** *$\mathbb{P}(\mathcal{E}_2^c) \leq \mathbb{T}_1$, where $a_2 = \frac{64 b_z^2}{\sigma_z^2 p_z^2}$ and $a_3 = \frac{p_z^2}{8}$.*

Lemma 4.2 (Bound on $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2)$)**.** *$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{T}_2(\delta)$, where $a_4 = \max(1, 4 b_z \sqrt{n_x}/a_1)$.*

Roughly, Lemma 4.1 indicates that PE holds with high probability, which is proved by leveraging the BMSB assumption and set discretization. The proof of Lemma 4.2 is more involved and is our major technical contribution. On a high level, the proof relies on two technical lemmas below.

Lemma 4.3 (Discretization of $\mathcal{E}_1 \cap \mathcal{E}_2$ (informal))**.** *Let $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\}$ denote an ϵ_γ-net of $\{\gamma : \|\gamma\|_F = 1\}$. Under a proper choice of ϵ_γ, we have $v_\gamma = \tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}$.*[7] *We can construct $\tilde{\Gamma}_T$ such that*

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\{\exists\, 1 \leq i \leq v_\gamma, d \geq 0, \text{ s.t. } d\gamma_i \in \tilde{\Gamma}_T\} \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2)$$

where $\mathcal{E}_{1,i} = \{\exists\, d \geq 0, \text{ s.t. } d\gamma_i \in \tilde{\Gamma}_T\}$.

Lemma 4.3 leverages finite set discretization to bound the existence of a feasible element in an infinite continuous set. The formal version of Lemma 4.3 is provided as Lemma D.8 in the appendix.

Lemma 4.4 (Construction of event $G_{i,k}$ via stopping times (informal))**.** *Consider $\mathcal{F}_t$ as defined in Assumption 2.5. Under the conditions in Lemma 4.3, we construct $G_{i,k}$ for all i and all $0 \leq k \leq T/m - 1$ by*

$$G_{i,k} = \left\{ b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} \geq \frac{a_1 \delta}{4\sqrt{n_x}} - w_{\max}, \text{ and } \frac{1}{m} \sum_{s=1}^{m} z_{km+s} z_{km+s}^\top \succeq a_1^2 I_{n_z} \right\}.$$

where $b_{i,t}, j_{i,t}$ are measurable in $\mathcal{F}_t$, and $L_{i,k}$ is constructed as a stopping time with respect to $\{\mathcal{F}_{km+l}\}_{l \geq 0}$. The formal definitions of $b_{i,t}, j_{i,t}, L_{i,k}$ are provided in Appendix D.3.1.

Then, we have

$$\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}\left(\bigcap_{k=0}^{T/m-1} G_{i,k} \right) \leq \left(1 - q_w \left(\frac{a_1 \delta}{4\sqrt{n_x}} \right) \right)^{\frac{T}{m}}$$

The constructions of $G_{i,k}$ and $L_{i,k}$ in Lemma 4.4 are our major technical contribution. With the constructions above, the proof can be completed by leveraging the conditional independence property of stopping times, which is briefly discussed below. Notice that by conditioning on the event $\{L_{i,k} = l\}$, we have $w_{km+L_{i,k}} = w_{km+l}$ and w_{km+l} is independent of $\mathcal{F}_{km+l}$. Consequently, w_{km+l} is also independent of $b_{i,km+L_{i,k}}, j_{i,km+L_{i,k}}$ conditioning on $\{L_{i,k} = l\}$ since $b_{i,km+l}, j_{i,km+l}$ are measurable in $\mathcal{F}_{km+l}$. Therefore, the probability of $G_{i,k}$ conditioning on $\{L_{i,k} = l\}$ can be bounded by the probability distribution of w_t, which enjoys good properties such as Assumption 2.6. More details of the proof are in Appendix D.3.3.

In conclusion, Lemma 4.2 follows directly from Lemma 4.3 and Lemma 4.4. Combining Lemma 4.2 and Lemma 4.1 completes the proof of Theorem 3.1.

Remark 4.5 (Convergence rate of SME for general time series)**.** *Similar to Theorem 2.4 in (Simchowitz et al., 2018), our results for linear dynamical systems can also be generalized to general time series with linear responses:*

$$y_t = \theta^* z_t + w_t, \quad t \geq 0,$$

where $\mathcal{F}_t^y = \mathcal{F}(w_0, \ldots, w_t, z_0, \ldots, z_t)$, $y_t \in \mathbb{R}^{n_y}$ is measurable in $\mathcal{F}_t^y$ but not in $\mathcal{F}_{t-1}^y$. The SME algorithm is

$$\Theta_T^y = \bigcap_{t=0}^{T-1} \{\hat{\theta} : y_t - \hat{\theta} z_t \in \mathcal{W}\}.$$

Under Assumptions 2.2, 2.5, and 2.6, we have

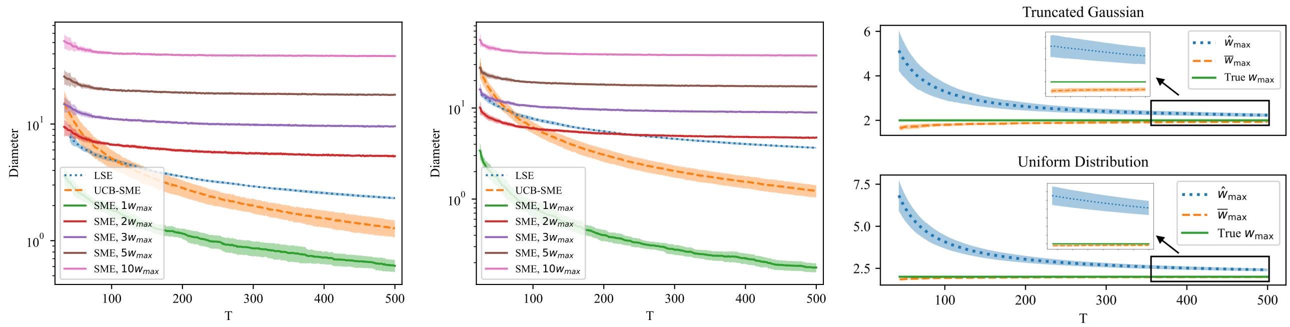
$$\mathbb{P}(\text{diam}(\Theta_T^y) > \delta) \leq \frac{T}{m} \tilde{O}(n_z^{2.5}) a_2^{n_z} \exp(-a_3 m) + \tilde{O}((n_y n_z)^{2.5}) a_4^{n_y n_z} \left(1 - q_w \left(\frac{a_1 \delta}{4\sqrt{n_y}} \right) \right)^{\lceil T/m \rceil},$$

where a_1, a_2, a_3 are defined in Theorem 3.1 and $a_4 = \frac{4 b_z \sqrt{n_y}}{a_1}$.

5. Applications to Robust Adaptive Control

Robust adaptive control usually involves two steps: updating the uncertainty set estimation, and designing robust

[7] The exact formulas of v_γ and ϵ_γ are in Lemma D.3.



(a) Truncated Gaussian (b) Uniform Distribution (c) $\hat{w}_{\max}^{(T)}$ in UCB-SME and estimation $\bar{w}_{\max}^{(T)}$

Figure 2. **Figures (a)-(b)** compares the diameters of SME, UCB-SME, and SME with loose disturbance upper bounds that are 2, 3, 5, and 10 times larger than the true disturbance bound $w_{\max}$, as well as the baseline uncertainty set from the 90% confidence region of LSE. **Figure (c)** shows the convergence to the true bound $w_{\max}$ of the lower estimation $\bar{w}_{\max}$ in (4) and the UCB $\hat{w}_{\max}$ generated by the UCB-SME algorithm in **Figures (a)-(b)**.

controllers based on the updated uncertainty set. SME can be naturally applied to robust adaptive control as the updating rule of the uncertainty set estimation. To illustrate this, we discuss the applications of SME to two popular controllers, robust adaptive MPC and robust SLS. We focus on the implications of our convergence rates.

Application of SME to robust adaptive MPC. SME has long been adopted in the robust adaptive MPC design (see e.g., (Köhler et al., 2019; Lorenzen et al., 2019; Lu & Cannon, 2023)). However, due to the lack of non-asymptotic guarantees for SME, the non-asymptotic analysis for robust adaptive MPC also remains unsolved. Applying Theorem 3.1 straightforwardly, we can obtain a non-asymptotic estimation error bound for robust adaptive MPC below, which lays a foundation for future regret analysis. For simplicity, we consider a tight bound $\mathcal{W}$ is known below, but our results for unknown $\mathcal{W}$ can also be applied similarly.

Corollary 5.1. *Consider the robust adaptive MPC controller introduced in Example 1, where Θ_t is updated by SME and $\mathcal{W}$ is known.*[8] *Under the conditions of Corollary 3.2, the estimation error for any $\hat{\theta}_T \in \Theta_T$ can be bounded by $\|\hat{\theta}_T - \theta^*\|_F \leq \tilde{O}(\frac{n_x^{1.5} n_z^2}{T})$ with high probability.*

Application of SME to robust SLS. Robust SLS has been proposed in (Dean et al., 2019b) for robust constrained control under system uncertainties (Dean et al., 2019b). Since (Dean et al., 2019b) assumes bounded disturbances, one can apply SME for the uncertainty set estimation in place of the LSE's confidence regions in (Dean et al., 2019b). Then, by leveraging Theorems 3.1, 4.1 in (Dean et al., 2019b) and our Theorem 3.1, we can directly obtain a non-asymptotic sub-optimality gap for learning-based robust SLS with SME as the uncertainty set estimation. For simplicity, we consider a known tight bound $\mathcal{W}$, but our results for unknown $\mathcal{W}$ can also be similarly applied here.

Corollary 5.2. *Under the conditions in Theorem 3.1 in (Dean et al., 2019b) and Corollary 3.2, for large enough T, we have $\frac{J(A^*,B^*,\hat{\mathbf{K}})-J^*}{J^*} \leq \tilde{O}(n_x^{1.5} n_z^2/T)$, where $\hat{\mathbf{K}}$ denotes the robust SLS controller in (Dean et al., 2019b) under the uncertainty set Θ_T constructed by SME, $J(A^*, B^*, \hat{\mathbf{K}}) = \lim_{T\to+\infty} \mathbb{E} \frac{1}{T} \sum_{t=0}^{T-1} (x_t^\top Q x_t + u_t^\top R u_t)$ denotes the infinite-horizon averaged total cost by implementing the robust SLS controller $\hat{\mathbf{K}}$, and J^* denotes the optimal infinite-horizon averaged total cost.*

6. Numerical Experiments

We evaluate the empirical performance of SME on various systems and applications. For all experiments, we use the 90% confidence regions of LSE computed by Lemma E.3 in (Simchowitz & Foster, 2020) and Theorem 1 in (Abbasi-Yadkori & Szepesvári, 2011) as the baseline. The details of the simulation settings are provided in Appendix I.[9]

Comparison of SME, SME with loose bound, UCB-SME, and LSE. This experiment is based on the linearized longitudinal flight control dynamics of Boeing 747 as studied in recent literature on learning-based control of linear systems (Lale et al., 2022; Mete et al., 2022).

In Figure 2, we show the diameters of SME, SME with loose disturbance bounds, and UCB-SME on the identification problem of the Boeing 747 dynamics with i.i.d. truncated Gaussian (Figure 2a) and uniform (Figure 2b) disturbances. We use control actions sampled from a uniform distribution in both cases. In Figure 2c, we show that both the upper bound $\hat{w}_{\max}$ used for UCB-SME and the lower bound $\bar{w}_{\max}$

[8] When $\mathcal{W}$ is unknown, Theorems 3.4-3.7 all apply.

[9] The code to reproduce all the experimental results can be found at https://github.com/jy-cds/non-asymptotic-set-membership.

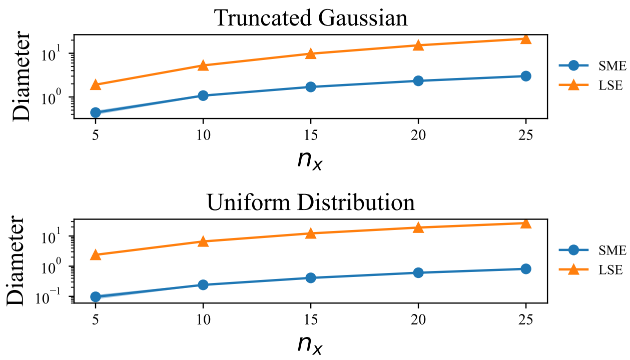


Figure 3. Diameters of the uncertainty sets constructed by SME, UCB-SME, and LSE for systems with different dimensions.

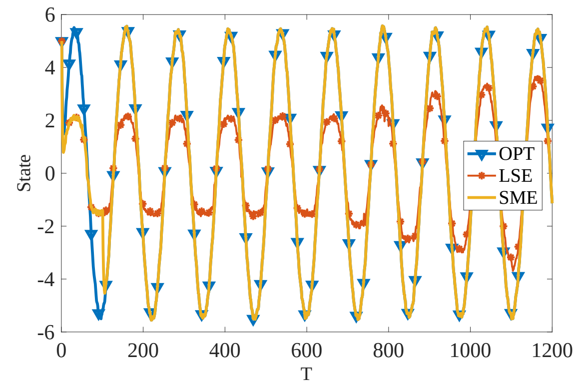


Figure 4. Linear quadratic tracking of robust adaptive MPC based on SME, LSE's confidence regions, and the accurate model (OPT).

in (4) converge to the true bound $w_{\max}$ as T increases. The quantitative behaviors of SME and its variants are consistent with those predicted by our theoretical results. In particular, in Figure 2a and Figure 2b, SME and UCB-SME outperform the 90% confidence regions of LSE in both the magnitude and the convergence rate. In Figure 2c, we verify that the UCB estimation $\hat{w}_{\max}^{(T)}$ converges to the true disturbance bound $w_{\max}$ from above, while the estimation $\bar{w}_{\max}^{(T)}$ converges from below. It is worth noting that $\bar{w}_{\max}^{(T)}$ converges to $w_{\max}$ very quickly in the simulations, allowing $\bar{w}_{\max}^{(T)}$ to be another potential approximation of $w_{\max}$ for SME when T is very large.

Scaling with dimension. We compare the scaling of SME, SME-UCB, and LSE with respect to the system dimensions in Figure 3. We use an autonomous system $x_{t+1} = A^{\star}x_t + w_t$, where $A \in \mathbb{R}^{n_x \times n_x}$ has varying n_x. Disturbances w_t are sampled from a truncated Gaussian distribution and uniform distribution with $w_{\max} = 2$. Surprisingly, the scaling of SME with respect to the dimension of the system is not significantly worse than that of LSE in the simulation. This suggests that the convergence rate in Corollary 3.2 can potentially be improved in terms of the dimension dependence, which is left for future investigation.

Application to robust adaptive MPC. We provide an example of the quantitative impact of using SME for adaptive robust MPC in Figure 4. We consider the task of constrained linear quadratic tracking problem as in (Li et al., 2023). The model uncertainty set is estimated online with SME and LSE's 90 % confidence region. Control actions are computed using the tube-based robust MPC (Rawlings et al., 2017; Mayne et al., 2005) with the uncertainty sets. We also plot the optimal MPC controller with accurate model information. Thanks to the fast convergence of SME, the tracking performance of the tube-based robust MPC with SME estimation quickly coincides with OPT, while the same controller based on LSE's confidence region estimation converges more slowly.

7. Concluding Remarks

This work provides the first convergence rates for SME in linear dynamical systems with bounded disturbances and discusses variants of SME with unknown bound on w_t. Numerical experiments demonstrate SME's promising performance under bounded disturbances.

Regarding limitations and future directions, this work only considers box constraints on w_t, so it is worth extending the analysis to more general constraints. In addition, this paper only measures the size of the uncertainty sets by their diameters. We leave for future work to consider other metrics, such as volume. Further, our bounds suffer poor dependence on the system dimension, which is not reflected in simulations. Hence, it is important to further refine the bounds and discuss the fundamental limits. Another exciting direction is to speed up the computation of SME since the current computation complexity increases linearly with the sample size. The convergence rate of fast SME algorithms is an important *open question*. Other interesting directions include the extensions of the SME analysis to nonlinear systems, where recent nonlinear system identification literature (Sattar et al., 2022; Foster et al., 2020) may provide insights; and analyzing SME in the presence of other uncertainties, e.g. measurement noises (Sarkar & Rakhlin, 2019).

SME is a valid estimation for bounded non-stochastic disturbances (Fogel & Huang, 1982; Milanese et al., 2013; Lauricella & Fagiano, 2020; Livstone & Dahleh, 1996). Therefore, a fruitful direction is to study SME's convergence rates under non-stochastic w_t. Another potential method for uncertainty set estimation is the credible regions of Bayesian approaches, e.g. Thompson sampling for linear systems (Kargin et al., 2022; Abeille & Lazaric, 2017) and Gaussian processes for nonlinear systems (Fisac et al., 2018). A future direction is to study the convergence rates of credible regions.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

References

Abbasi-Yadkori, Y. and Szepesvári, C. Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th Annual Conference on Learning Theory*, pp. 1–26. JMLR Workshop and Conference Proceedings, 2011.

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. Online least squares estimation with self-normalized processes: An application to bandit problems. *arXiv preprint arXiv:1102.2670*, 2011.

Abeille, M. and Lazaric, A. Thompson sampling for linear-quadratic control problems. In *Artificial intelligence and statistics*, pp. 1246–1254. PMLR, 2017.

Adetola, V. and Guay, M. Robust adaptive mpc for constrained uncertain nonlinear systems. *International Journal of Adaptive Control and Signal Processing*, 25(2): 155–167, 2011.

Akçay, H. The size of the membership-set in a probabilistic framework. *Automatica*, 40(2):253–260, 2004.

Bai, E.-W., Tempo, R., and Cho, H. Membership set estimators: size, optimal inputs, complexity and relations with least squares. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 42(5):266–277, 1995.

Bai, E.-W., Cho, H., and Tempo, R. Convergence properties of the membership set. *Automatica*, 34(10):1245–1249, 1998.

Benevides, J. R., Paiva, M. A., Simplício, P. V., Inoue, R. S., and Terra, M. H. Disturbance observer-based robust control of a quadrotor subject to parametric uncertainties and wind disturbance. *IEEE Access*, 10:7554–7565, 2022.

Bertsekas, D. P. *Control of uncertain systems with a set-membership description of the uncertainty.* PhD thesis, Massachusetts Institute of Technology, 1971.

Boyd, S. and Vandenberghe, L. *Convex optimization*. Cambridge university press, 2004.

Brunke, L., Greeff, M., Hall, A. W., Yuan, Z., Zhou, S., Panerati, J., and Schoellig, A. P. Safe learning in robotics: From learning-based control to safe reinforcement learning. *Annual Review of Control, Robotics, and Autonomous Systems*, 5:411–444, 2022.

Bujarbaruah, M., Zhang, X., Tanaskovic, M., and Borrelli, F. Adaptive stochastic mpc under time-varying uncertainty. *IEEE Transactions on Automatic Control*, 66(6):2840–2845, 2020.

Campi, M. C. and Weyer, E. Finite sample properties of system identification methods. *IEEE Transactions on Automatic Control*, 47(8):1329–1334, 2002.

Chang, T.-J. and Shahrampour, S. Regret analysis of distributed online lqr control for unknown LTI systems. *IEEE Transactions on Automatic Control*, 69(1):667–673, 2024.

Chen, X. and Hazan, E. Black-box control for linear dynamical systems. In *Conference on Learning Theory*, pp. 1114–1143. PMLR, 2021.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. Regret bounds for robust adaptive control of the linear quadratic regulator. *Advances in Neural Information Processing Systems*, 31:4188–4197, 2018.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, pp. 1–47, 2019a.

Dean, S., Tu, S., Matni, N., and Recht, B. Safely learning to control the constrained linear quadratic regulator. In *2019 American Control Conference (ACC)*, pp. 5582–5588. IEEE, 2019b.

Eising, J. and Cortes, J. When sampling works in data-driven control: Informativity for stabilization in continuous time. *arXiv preprint arXiv:2301.10873*, 2023.

Eising, J., Liu, S., Martínez, S., and Cortés, J. Using data informativity for online stabilization of unknown switched linear systems. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 8–13. IEEE, 2022.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. Finite-time adaptive stabilization of linear systems. *IEEE Transactions on Automatic Control*, 64(8):3498–3505, 2018.

Fisac, J. F., Akametalu, A. K., Zeilinger, M. N., Kaynama, S., Gillula, J., and Tomlin, C. J. A general safety framework for learning-based control in uncertain robotic systems. *IEEE Transactions on Automatic Control*, 64(7): 2737–2752, 2018.

Fogel, E. and Huang, Y.-F. On the value of information in system identification—bounded noise case. *Automatica*, 18(2):229–238, 1982.

Foster, D., Sarkar, T., and Rakhlin, A. Learning nonlinear dynamical systems from a single trajectory. In *Learning for Dynamics and Control*, pp. 851–861. PMLR, 2020.

Hespanha, J. P. *Linear systems theory*. Princeton university press, 2018.

Hewing, L., Wabersich, K. P., Menner, M., and Zeilinger, M. N. Learning-based model predictive control: Toward safe learning in control. *Annual Review of Control, Robotics, and Autonomous Systems*, 3:269–296, 2020.

Ho, D., Le, H., Doyle, J., and Yue, Y. Online robust control of nonlinear systems with large uncertainty. In *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 130, pp. 3475–3483. PMLR, 4 2021.

Hsu, D., Kakade, S. M., and Zhang, T. Random design analysis of ridge regression. In *Conference on learning theory*, pp. 9–1. JMLR Workshop and Conference Proceedings, 2012.

Kargin, T., Lale, S., Azizzadenesheli, K., Anandkumar, A., and Hassibi, B. Thompson sampling achieves $\tilde{o}\sqrt{t}$ regret in linear quadratic control. In *Conference on Learning Theory*, pp. 3235–3284. PMLR, 2022.

Kitamura, W. and Fujisaki, Y. Convergence properties of the membership set in the presence of disturbance and parameter uncertainty. *Transactions of the Society of Instrument and Control Engineers*, 39(4):382–387, 2003.

Kitamura, W., Fujisaki, Y., and Bai, E.-W. The size of the membership set in the presence of disturbance and parameter uncertainty. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 5698–5703. IEEE, 2005.

Köhler, J., Andina, E., Soloperto, R., Müller, M. A., and Allgöwer, F. Linear robust adaptive model predictive control: Computational complexity and conservatism. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 1383–1388. IEEE, 2019.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Reinforcement learning with fast stabilization in linear dynamical systems. In *International Conference on Artificial Intelligence and Statistics*, pp. 5354–5390. PMLR, 2022.

Lauricella, M. and Fagiano, L. Set membership identification of linear systems with guaranteed simulation accuracy. *IEEE Transactions on Automatic Control*, 65(12): 5189–5204, 2020.

Li, Y., Das, S., and Li, N. Online optimal control with affine constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 8527–8537, 2021a.

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

Li, Y., Zhang, T., Das, S., Shamma, J., and Li, N. Non-asymptotic system identification for linear systems with nonlinear policies. *arXiv preprint arXiv:2306.10369*, 2023.

Liu, X., Yang, Z., and Ying, L. Online nonstochastic control with adversarial and static constraints. *arXiv preprint arXiv:2302.02426*, 2023.

Liu, Y., Zhao, Y., and Wu, F. Ellipsoidal state-bounding-based set-membership estimation for linear system with unknown-but-bounded disturbances. *IET Control Theory & Applications*, 10(4):431–442, 2016.

Livstone, M. M. and Dahleh, M. A. Asymptotic properties of set membership identification algorithms. *Systems & control letters*, 27(3):145–155, 1996.

Ljung, L. *System identification*. Springer, 1998.

Lopez, B. T., Slotine, J.-J. E., and How, J. P. Robust adaptive control barrier functions: An adaptive and data-driven approach to safety. *IEEE Control Systems Letters*, 5(3): 1031–1036, 2020.

Lorenzen, M., Cannon, M., and Allgöwer, F. Robust mpc with recursive model update. *Automatica*, 103:461–471, 2019.

Lu, X. and Cannon, M. Robust adaptive model predictive control with persistent excitation conditions. *Automatica*, 152:110959, 2023.

Lu, X., Cannon, M., and Koksal-Rivet, D. Robust adaptive model predictive control: Performance and parameter estimation. *International Journal of Robust and Nonlinear Control*, 2019.

Mania, H., Tu, S., and Recht, B. Certainty equivalence is efficient for linear quadratic control. In *Advances in Neural Information Processing Systems*, volume 32, pp. 10154–10164. Curran Associates, Inc., 2019.

Mayne, D. Q., Seron, M. M., and Rakovic, S. Robust model predictive control of constrained linear systems with bounded disturbances. *Automatica*, 41(2):219–224, 2005.

Mete, A., Singh, R., and Kumar, P. Augmented rbmle-ucb approach for adaptive control of linear quadratic systems. *Advances in Neural Information Processing Systems*, 35: 9302–9314, 2022.

Milanese, M. and Vicino, A. Optimal estimation theory for dynamic systems with set membership uncertainty: An overview. *Automatica*, 27(6):997–1009, 1991.

Milanese, M., Norton, J., Piet-Lahanier, H., and Walter, É. *Bounding approaches to system identification*. Springer Science & Business Media, 2013.

Narendra, K. and Annaswamy, A. Robust adaptive control in the presence of bounded disturbances. *IEEE Transactions on Automatic Control*, 31(4):306–315, 1986.

Narendra, K. S. and Annaswamy, A. M. Persistent excitation in adaptive systems. *International Journal of Control*, 45 (1):127–160, 1987.

Oymak, S. and Ozay, N. Non-asymptotic identification of lti systems from a single trajectory. In *2019 American control conference (ACC)*, pp. 5655–5661. IEEE, 2019.

Parsi, A., Iannelli, A., and Smith, R. S. Active exploration in adaptive model predictive control. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pp. 6186–6191. IEEE, 2020a.

Parsi, A., Iannelli, A., Yin, M., Khosravi, M., and Smith, R. S. Robust adaptive model predictive control with worst-case cost. *IFAC-PapersOnLine*, 53(2):4222–4227, 2020b.

Petrik, M. and Russel, R. H. Beyond confidence regions: Tight bayesian ambiguity sets for robust mdps. *Advances in neural information processing systems*, 32, 2019.

Qi, J., He, G., Mei, S., and Liu, F. Power system set membership state estimation. In *2012 IEEE Power and Energy Society General Meeting*, pp. 1–7. IEEE, 2012.

Rantzer, A. Concentration bounds for single parameter adaptive control. In *2018 Annual American Control Conference (ACC)*, pp. 1862–1866. IEEE, 2018.

Rawlings, J. B. and Mayne, D. Q. *Model predictive control: Theory and design*. Nob Hill Pub., 2009.

Rawlings, J. B., Mayne, D. Q., and Diehl, M. *Model predictive control: theory, computation, and design*, volume 2. Nob Hill Publishing Madison, WI, 2017.

Rogers, C. Covering a sphere with spheres. *Mathematika*, 10(2):157–164, 1963.

Sarkar, T. and Rakhlin, A. Near optimal finite time identification of arbitrary linear dynamical systems. In *International Conference on Machine Learning*, pp. 5610–5618. PMLR, 2019.

Sarker, A., Fisher, P., Gaudio, J. E., and Annaswamy, A. M. Accurate parameter estimation for safety-critical systems with unmodeled dynamics. *Artificial Intelligence*, pp. 103857, 2023.

Sasfi, A., Zeilinger, M. N., and Köhler, J. Robust adaptive mpc using control contraction metrics. *arXiv preprint arXiv:2209.11713*, 2022.

Sattar, Y., Oymak, S., and Ozay, N. Finite sample identification of bilinear dynamical systems. In *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 6705–6711. IEEE, 2022.

Simchowitz, M. and Foster, D. Naive exploration is optimal for online lqr. In *International Conference on Machine Learning*, pp. 8937–8948. PMLR, 2020.

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

Tanaskovic, M., Fagiano, L., Smith, R., Goulart, P., and Morari, M. Adaptive model predictive control for constrained linear systems. In *2013 European Control Conference (ECC)*, pp. 382–387. IEEE, 2013.

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

Tu, S. L. *Sample complexity bounds for the linear quadratic regulator*. University of California, Berkeley, 2019.

Van Overschee, P. and De Moor, B. *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 2012.

van Waarde, H. J., Eising, J., Camlibel, M. K., and Trentelman, H. L. The informativity approach to data-driven analysis and control. *arXiv preprint arXiv:2302.10488*, 2023.

Verger-Gaugry, J.-L. Covering a ball with smaller equal balls in $\mathbb{R}^n$. *Discrete and Computational Geometry*, 33: 143–155, 2005.

Vidyasagar, M. and Karandikar, R. L. A learning theory approach to system identification and stochastic adaptive control. *Probabilistic and randomized methods for design under uncertainty*, pp. 265–302, 2006.

Wabersich, K. P., Taylor, A. J., Choi, J. J., Sreenath, K., Tomlin, C. J., Ames, A. D., and Zeilinger, M. N. Data-driven safety filters: Hamilton-jacobi reachability, control barrier functions, and predictive methods for uncertain systems. *IEEE Control Systems Magazine*, 43(5):137–177, 2023.

Wagenmaker, A. and Jamieson, K. Active learning for identification of linear dynamical systems. In *Conference on Learning Theory*, pp. 3487–3582. PMLR, 2020.

Wu, C., Teo, K. L., and Wu, S. Min–max optimal control of linear systems with uncertainty and terminal state constraints. *Automatica*, 49(6):1809–1815, 2013.

Yeh, C., Yu, J., Shi, Y., and Wierman, A. Robust online voltage control with an unknown grid topology. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, pp. 240–250, 2022.

Yu, J., Gupta, V., and Wierman, A. Online stabilization of unknown linear time-varying systems. *arXiv preprint arXiv:2304.02878*, 2023a.

Yu, J., Ho, D., and Wierman, A. Online adversarial stabilization of unknown networked systems. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 7(1):1–43, 2023b.

Zhang, K., Sun, Q., and Shi, Y. Trajectory tracking control of autonomous ground vehicles using adaptive learning mpc. *IEEE Transactions on Neural Networks and Learning Systems*, 32(12):5554–5564, 2021.

Zhang, X., Shi, W., Li, X., Yan, B., Malkawi, A., and Li, N. Decentralized temperature control via HVAC systems in energy efficient buildings: An approximate solution procedure. In *Proceedings of 2016 IEEE Global Conference on Signal and Information Processing*, pp. 936–940, 2016.

Zhao, Z. and Li, Q. Adaptive sampling methods for learning dynamical systems. In *Mathematical and Scientific Machine Learning*, pp. 335–350. PMLR, 2022.

Zheng, Y. and Li, N. Non-asymptotic identification of linear dynamical systems using multiple trajectories. *IEEE Control Systems Letters*, 5(5):1693–1698, 2020.

Roadmap for the appendices

- Appendix A introduces additional notation used throughout the Appendix.
- Appendix B provides more literature review on LSE and SM, and a more detailed discussion on the technical contributions of this paper.
- Appendix C provides more discussions on examples that satisfy Assumptions 2.5 and 2.6.
- Appendix D presents the proof of Theorem 3.1. In particular, we provide helper lemmas in Appendix D.1 and prove Lemma 4.1, Lemma 4.2 in Appendix D.2 and Appendix D.3 respectively. A more precise upper bound for Theorem 3.1 (without the $\tilde{O}(\cdot)$ notation) is provided in Appendix D.4.
- Appendix E presents a proof of Corollary 3.2
- Appendix F provides a proof of Corollary 3.3.
- Appendix G presents a proof of Theorem 3.4.
- Appendix H provides proofs of Theorem 3, Corollary 3, and Theorem 4.
- Appendix I provides details of the simulation.

A. Additional notations

Let $\mathbb{S}_n(0,1)$ denote the unit sphere in $\mathbb{R}^n$ in l_2 norm, i.e., $\mathbb{S}_n(0,1)=\{x\in\mathbb{R}^n : \|x\|_2=1\}$. Let $\mathbb{S}_{n\times m}(0,1)$ denote the unit sphere in $\mathbb{R}^{n\times m}$ with respect to the Frobenius norm, i.e., $\mathbb{S}_{n\times m}(0,1)=\{M\in\mathbb{R}^{n\times m} : \|M\|_F=1\}$. Let $\bar{B}_n(0,1)$ denote the closed unit ball in $\mathbb{R}^n$ in l_2 norm, i.e., $\bar{B}_n(0,1)=\{x\in\mathbb{R}^n : \|x\|_2\leq 1\}$. Let $\bar{B}_{n\times m}(0,1)$ denote the closed unit ball in $\mathbb{R}^{n\times m}$ in Frobenius norm, i.e., $\bar{B}_{n\times m}(0,1)=\{M\in\mathbb{R}^{n\times m} : \|M\|_F\leq 1\}$. For a matrix $M\in\mathbb{R}^{n\times m}$, vec$(M)$ is the vectorization of M. Moreover, we define the inverse mapping of vec$(\cdot)$ as mat$(\cdot)$, i.e., for a vector $d\in\mathbb{R}^{nm}$, mat$(d)\in\mathbb{R}^{n\times m}$. Consider a σ-algebra $\mathcal{F}$ and a random variable X, we write $X\in\mathcal{F}$ if X is measurable with respect to $\mathcal{F}$, i.e., for all Borel measurable sets $B\subseteq\mathbb{R}$, we have $X^{-1}(B)\in\mathcal{F}$. We can similarly define $\mathcal{F}$-measurable random matrices and random vectors. Further, consider a polyhedral $\mathbb{D}=\{x : Ax\leq b\}$, we write $\mathbb{D}\in\mathcal{F}$ if matrix A and vector b are measurable with respect to $\mathcal{F}$. Consider two symmetric matrices $A,B\in\mathbb{R}^{n\times n}$, we write $A\succeq B$ if $A-B$ is a positive definite matrix. We define $\min\emptyset=+\infty$. For a set $\mathcal{E}$, let $\mathbb{1}_{\mathcal{E}}$ denote the indicator function on $\mathcal{E}$. For a vector $x\in\mathbb{R}^n$, we use x^j to denote the jth coordinate of x. Throughout the paper, we use TrunGauss$(0,\sigma_w,[-w_{\max},w_{\max}])$ to refer to the truncated Gaussian distribution generated by Gaussian distribution with zero mean and σ_w^2 variance with truncated range $[-w_{\max},w_{\max}]$. The same applies to multi-variate truncated Gaussian distributions.

B. More discussions on least square and set membership

System identification studies the problem of estimating the parameters of an unknown dynamical systems from trajectory data. There are two main classes of estimation methods: point estimator such as least square estimation (LSE), and set estimator such as set membership estimation (SME). In the following, we provide more discussions and literature review on LSE and SME. We will also discuss the major technical novelties of this work.

B.1. Least square estimation

For linear dynamical systems $x_{t+1}=A^*x_t+B^*u_t+w_t=\theta^*z_t+w_t$, given a trajectory of data $\{x_t,u_t\}_{t\geq 0}$, least square estimation generates a point estimator that minimizes the following quadratic error (Van Overschee & De Moor, 2012; Ljung, 1998):

$$\hat{\theta}_{\text{LSE}}=\min_{\hat{\theta}}\sum_{t=0}^{T-1}\|x_{t+1}-\hat{\theta}z_t\|_2^2.$$

Least-square estimation is widely used and its convergence (rate) guarantees have been investigated for a long time. In particular, non-asymptotic convergence rate guarantees of LSE has become increasingly important as these guarantees are the foundations for non-asymptotic performance analysis of learning-based/adaptive control algorithms. Earlier non-asymptotic analysis of LSE focused on the simpler regression model $y_t = \theta^* x_t + w_t$, where x_t and y_t are independent (Campi & Weyer, 2002; Vidyasagar & Karandikar, 2006; Hsu et al., 2012).

Recently, there is one major breakthrough in (Simchowitz et al., 2018) that provides LSE's convergence rate analysis for linear dynamical system $x_{t+1} = \theta^* z_t + w_t$, where x_{t+1} and $z_t = [x_t^\top, u_t^\top]^\top$ are correlated. More specifically, (Simchowitz et al., 2018) establishes a fundamental property, *block-martingale small-ball (BMSB)*, to analyze LSE under correlated data. BMSB enables a long list of subsequent literature on LSE's non-asymptotic analysis for different types of dynamical systems, e.g., (Oymak & Ozay, 2019; Dean et al., 2019a; Zheng & Li, 2020; Rantzer, 2018; Faradonbeh et al., 2018; Wagenmaker & Jamieson, 2020; Tsiamis et al., 2022; Zhao & Li, 2022; Li et al., 2021b).

Though LSE is a point estimator, one can establish confidence region of LSE based on proper statistical assumptions on w_t. The pioneer works on the confidence region of LSE for linear dynamical systems are (Abbasi-Yadkori et al., 2011; Abbasi-Yadkori & Szepesvári, 2011), which construct ellipsoid confidence regions for LSE. Moreover, the non-asymptotic bounds on estimation errors established in (Simchowitz et al., 2018; Dean et al., 2019b) can also be viewed as confidence bounds. Further, the estimation error $\tilde{O}(\frac{\sqrt{n_x+n_z}}{\sqrt{T}})$ has been shown to match the fundamental lower bound for any estimation methods for unbounded disturbances in (Simchowitz et al., 2018). However, these confidence bounds all rely on statistical inequalities, which may result in loose constant factors despite an optimal convergence rate. When applying these confidence bounds to robust control, where the controller is required to satisfy certain stability and constraint satisfaction properties for every possible system in the confidence region, a loose constant factor will result in a larger confidence region and a more conservative control design. Finally, in robust control and many practical applications, the disturbances are usually bounded, and it will be interesting to see how the knowledge of the boundedness will improve the uncertainty set estimation.

On a side note, this paper is also related with the ambiguity set estimation for the transition probabilities in robust Markov decision processes (Petrik & Russel, 2019). There are attempts on improving the ambiguity set estimation based on LSE for less conservative robust MDP (Petrik & Russel, 2019).

B.2. Set membership

Set membership is commonly used in robust control for uncertainty set estimation (Milanese & Vicino, 1991; Adetola & Guay, 2011; Tanaskovic et al., 2013; Bujarbaruah et al., 2020; Zhang et al., 2021; Parsi et al., 2020b;a; Sasfi et al., 2022). There is a long history of research on SME for both deterministic disturbances, such as (Bai et al., 1995; Fogel & Huang, 1982; Kitamura & Fujisaki, 2003; Milanese et al., 2013; Lauricella & Fagiano, 2020; Livstone & Dahleh, 1996), and stochastic disturbances, such as (Bai et al., 1998; 1995; Kitamura et al., 2005; Akçay, 2004; Lu et al., 2019). For the stochastic disturbances, both convergence and convergence rate analysis have been investigated under the persistent excitation (PE) condition. However, the existing convergence rates are only established for simpler regression problems, $y_t = \theta^* x_t + w_t$, where y_t and x_t are independent (Akçay, 2004; Bai et al., 1995; 1998; Kitamura et al., 2005).

Recently, (Lu et al., 2019) provided an initial attempt to establish the convergence guarantee of SME for linear dynamical systems $x_{t+1} = \theta^* z_t + w_t$ for correlated data x_{t+1} and z_t. However, (Lu et al., 2019) assumes that PE holds deterministically, and designs a special control design based on constrained optimization to satisfy PE deterministically. Therefore, the convergence for general control design and the convergence rate analysis remain open questions for correlated data arising from dynamical systems.

In this paper, we establish the convergence rate guarantees of SME on linear dynamical systems under the BMSB conditions in (Simchowitz et al., 2018). Compared with (Lu et al., 2019), BMSB condition can be satisfied by adding an i.i.d. random noise to a general class of control designs (Li et al., 2021b).

Technically, one major challenge of SME analysis compared with the LSE analysis is that the diameter of the membership set does not have an explicit formula, which is in stark contrast with LSE, where the point estimator is the solution to a quadratic program and has explicit form. A common trick to address this issue in the analysis of SME is to connect the diameter bound with the values of disturbances subsequences $\{w_{s_k}\}_{k\geq 0}$: it can be generally shown that a large diameter indicates that a long subsequence of disturbances are far away from the boundary of $\mathcal{W}$. However, existing construction methods of $\{w_{s_k}\}_{k\geq 0}$ will cause the time indices $\{s_k\}_{k\geq 0}$ to *correlate* with the realization of the sequences $\{x_t, u_t, w_t\}_{t\geq 0}$ (Akçay, 2004; Lu et al., 2019; Bai et al., 1995).[10] Consequently, in the correlated-data scenario and when PE does not hold

[10] In (Lu et al., 2019), the correlation between $\{s_k\}_{k\geq 0}$ and $\{x_t, u_t, w_t\}_{t\geq 0}$ is via the PE condition, but (Lu et al., 2019) assume deterministic PE to avoid this correlation issue.

deterministically, under the existing construction methods in (Akçay, 2004; Lu et al., 2019; Bai et al., 1995), the probability of $\{w_{s_k}\}_{k\geq 0}$ with correlated time indices *cannot* be bounded by the probability of the *independent* sequence $\{w_t\}_{t\geq 0}$. One major **technical contribution** of this paper is to provide a novel construction of $\{w_{s_k}\}_{k\geq 0}$ based on a sequence of stopping times and establish conditional independence properties despite correlated data and stochastic PE condition (BMSB). More details can be found in Lemma 4.4 and the proof or Lemma 4.2.

Though we only consider box constraints for w_t, it is worth mentioning that SME can be applied to much more general forms of disturbances. For example, a common alternative is the ellipsoidal-bounded disturbance where $\mathcal{W} := \{w \in \mathcal{R}^{n_x} : w^\top P w \leq 1\}$ with positive definite $P \in \mathcal{R}^{n_x \times n_x}$ (Bai et al., 1995; van Waarde et al., 2023; Eising & Cortes, 2023; Liu et al., 2016) and polytopic-bounded disturbance $\mathcal{W} := \{w \in \mathcal{R}^{n_x} : Gw \leq h\}$ for positive definite $G \in \mathcal{R}^{n_x \times n_x}$ and $h \in \mathcal{R}^{n_x}$ (Fogel & Huang, 1982; Lu et al., 2019; Lu & Cannon, 2023). There are also SME literature assuming bounded energy of the disturbance sequences (Bai et al., 1995). It is an interesting future direction to extend the analysis in this paper to more general disturbance constraints.

A separate but important challenge is that the knowledge of $\mathcal{W}$ is not always available a priori. There is literature discussing the estimation of $\mathcal{W}$ (Lauricella & Fagiano, 2020; Bai et al., 1998). We leave for future work how to simultaneously estimate $\mathcal{W}$ and perform non-asymptotic analysis on the size of the membership set.

Further, exact SME involves the intersection of an increasing number of sets, thus causing the computation complexity increases with time t, which can become prohibitive when t is large. There are many methods trying to reduce the computation complexity by approximating the membership sets (see e.g., (Livstone & Dahleh, 1996; Lu et al., 2019), etc.). It is an exciting future direction to study the diameter bounds of the approximated SME methods.

Lastly, it is worth pointing out that SME and its convergence rate in Theorem 3.1 can be easily extended to the general time series with linear responses below, which is also considered in LSE's convergence rate analysis in (Simchowitz et al., 2018). Same as (Simchowitz et al., 2018), we define the general time series with linear responses as $y_t = \theta^* z_t + w_t$, where $y_t, w_t \in \mathbb{R}^{n_y}$, $z_t \in \mathbb{R}^{n_z}$, $\theta^* \in \mathbb{R}^{n_y \times n_z}$. We let $\mathcal{F}_t = \sigma(w_0, \ldots, w_t, z_0, \ldots, z_t)$ be the natural filtration. Note that we consider $y_t \in \mathcal{F}_t$ but $y_t \notin \mathcal{F}_{t-1}$. It is straightforward to see that Theorem 3.1 still holds for this general time series since the proof in Appendix D does not require y_t to be the first n_x elements of z_{t+1}.

C. More discussions on Assumptions 2.5 and 2.6

C.1. More discussions on Assumption 2.5

The BMSB condition has been widely used in learning-based control. It has been shown that BMSB can be satisfied in many scenarios. For example, (Simchowitz et al., 2018; Tu, 2019) showed that linear systems with i.i.d. perturbed linear control policies, i.e., $x_{t+1} = Ax_t + B(Kx_t + \eta_t) + w_t$,[11] satisfy BMSB if the disturbances w_t and η_t are i.i.d. and follow Gaussian distributions with positive definite covariance matrices. Later, (Dean et al., 2019b) showed that $x_{t+1} = Ax_t + B(Kx_t + \eta_t) + w_t$ can still satisfy BMSB even for non-Gaussian distributions of w_t, η_t, as long as w_t and η_t have independent coordinates and finite fourth moments. Recently, (Li et al., 2021b) extended the results to linear systems with nonlinear policies, i.e., $x_{t+1} = Ax_t + B(\pi_t(x_t) + \eta_t) + w_t$, and showed that BMSB still holds as long as the nonlinear policies π_t generate bounded trajectories of states and control inputs, and w_t, η_t are bounded and follow distributions with certain anti-concentrated properties (a special case is positive definite covariance matrix).

C.2. More discussions on Assumption 2.6

In this subsection, we provide two example distributions, truncated Gaussian and uniform distributions, and discuss their corresponding $q_w(\epsilon)$ functions. It will be shown that for both distributions below, $q_w(\epsilon) = O(\epsilon)$.

Lemma C.1 (Example of uniform distribution). *Consider w_t that follows a uniform distribution on $[-w_{\max}, w_{\max}]^{n_x}$. Then, $q_w(\epsilon) = \frac{\epsilon}{2w_{\max}}$.*

Proof. Since Unif($\mathcal{W}$) is symmetric, we only need to consider one direction $j = 1$.

$$\mathbb{P}(w^j + w_{\max} \leq \epsilon) = \int_{w^1 + w_{\max} \leq \epsilon} \int_{w^2, \ldots, w^{n_x} \in [-w_{\max}, w_{\max}]} \frac{1}{(2w_{\max})^{n_x}} \mathbb{1}_{(w \in \mathcal{W})} \, \mathrm{d}w$$

[11]Though we only describe a static linear policy $u_t = Kx_t$ here, the results in (Simchowitz et al., 2018; Tu, 2019; Dean et al., 2019b) hold for dynamic linear policies.

$$= \int_{w^1 \le \epsilon - w_{\max}} \frac{1}{2w_{\max}} \mathbb{1}_{(w \in \mathcal{W})} \mathrm{d}w^1 = \frac{\epsilon}{2w_{\max}}$$

Similarly, $\mathbb{P}(w_{\max} - w^1 \le \epsilon) = \int_{w^1 \ge w_{\max} - \epsilon} \frac{1}{2w_{\max}} \mathbb{1}_{(w \in \mathcal{W})} dw^1 = \frac{\epsilon}{2w_{\max}}$. □

Lemma C.2 (Example of truncated Gaussian distribution)**.** *Consider w_t follows a truncated Gaussian distribution on $[-w_{\max}, w_{\max}]^{n_x}$ generated by a Gaussian distribution with zero mean and $\sigma_w I_{n_x}$ covariance matrix. Then,* $q_w(\epsilon) = \frac{1}{\min(\sqrt{2\pi}\sigma_w, 2w_{\max})} \exp(\frac{-w_{\max}^2}{2\sigma_w^2})\epsilon$.

Proof. Since this distribution is symmetric and each coordinate is independent, we only need to consider one direction j. Let X denote a Gaussian distribution with zero mean and σ_w^2 variance. By the definition of truncated Gaussian distributions, we have

$$\mathbb{P}(w^j + w_{\max} \le \epsilon) = \frac{\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon)}{\mathbb{P}(-w_{\max} \le X \le w_{\max})}$$

Notice that X/σ_w follows the standard Gaussian distribution, so we can obtain the following bounds.

$$\begin{aligned}\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon) &= \int_{-w_{\max}/\sigma_w}^{(-w_{\max}+\epsilon)/\sigma_w} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \,\mathrm{d}z \\ &\ge \frac{1}{\sqrt{2\pi}} \exp(-w_{\max}^2/(2\sigma_w^2)) \frac{\epsilon}{\sigma_w}\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(-w_{\max} \le X \le w_{\max}) &= \int_{-w_{\max}/\sigma_w}^{w_{\max}/\sigma_w} \frac{1}{\sqrt{2\pi}} \exp(-\frac{z^2}{2}) \,\mathrm{d}z \\ &\le \min(1, \frac{1}{\sqrt{2\pi}} \frac{2w_{\max}}{\sigma_w})\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathbb{P}(w^j + w_{\max} \le \epsilon) &= \frac{\mathbb{P}(-w_{\max} \le X \le -w_{\max} + \epsilon)}{\mathbb{P}(-w_{\max} \le X \le w_{\max})} \\ &\ge \max(\frac{1}{\sqrt{2\pi}} \exp(-w_{\max}^2/\sigma_w^2) \frac{\epsilon}{\sigma_w}, \frac{\epsilon}{2w_{\max}} \exp(\frac{-w_{\max}^2}{2\sigma_w^2})) \\ &= \frac{1}{\min(\sqrt{2\pi}\sigma_w, 2w_{\max})} \exp(\frac{-w_{\max}^2}{2\sigma_w^2})\epsilon\end{aligned}$$

Finally, $\mathbb{P}(w_{\max} - w^1 \le \epsilon)$ can be bounded similarly. □

Lemma C.3 (Example of uniform distribution on the boundary of $\mathcal{W}$ (a generalization of Rademacher distribution))**.** *Consider w_t follows a uniform distribution on $\{w : \|w\|_\infty = w_{\max}\}$. Then $q_w(\epsilon) = \frac{1}{2n_x}$.*

Proof. Since the hyper-cube $\{w : \|w\|_\infty = w_{\max}\}$ has $2n_x$ facets, the probability on each facet is $\frac{1}{2n_x}$. Therefore, $\mathbb{P}(w^j \le \epsilon - w_{\max}) \ge \mathbb{P}(w^j = -w_{\max}) = \frac{1}{2n_x}$ for all j. The same applies to $\mathbb{P}(w^j \ge -\epsilon + w_{\max})$. □

D. Proof of Theorem 3.1

The section provides more details for the proof of Theorem 3.1. In particular, we first provide technical lemmas for set discretization, then prove Lemma 4.1 and Lemma 4.2 respectively. The proof of Theorem 3.1 follows naturally by combining Lemma 4.1 and Lemma 4.2.

D.1. Technical lemmas: set discretization

This subsection provide useful technical lemmas for the proofs of Lemma 4.1 and Lemma 4.2. The results are based on a finite-ball covering result that is classical in the literature (Rogers, 1963) (Verger-Gaugry, 2005).

Theorem D.1 (Theorem 1.1 and 1.2 in (Verger-Gaugry, 2005) and Theorem 2 in (Rogers, 1963) (revised to match the setting of this paper))**.** *Consider a closed ball $\bar{\mathbb{B}}_n(0,1)=\{x\in\mathbb{R}^n : \|x\|_2\leq 1\}$ in l_2 norm. Considering covering this ball $\bar{\mathbb{B}}_n(0,1)$ with smaller closed balls $\bar{\mathbb{B}}_n(z,\epsilon)$ for $z\in\mathbb{R}^n$. Let $v_{\epsilon,n}$ denote the minimal number of smaller balls needed to cover $\bar{\mathbb{B}}_n(0,1)$. For $n\geq 1$ and $0<\epsilon<1/2$, we have*

$$v_{\epsilon,n}\leq 544n^{2.5}\log(n/\epsilon)(\frac{1}{\epsilon})^n$$

Proof. Theorem 1.1 and 1.2 in (Verger-Gaugry, 2005) and Theorem 2 in (Rogers, 1963) discuss the upper bounds of $v_{\epsilon,n}$ in several different cases. These upper bounds in these different cases are unified by the upper bound in the theorem above by algebraic manipulations. □

We apply Theorem D.1 to obtain the number of covering balls in the two settings below. These two settings will be considered in the proofs of Lemma 4.1 and 4.2 respectively.

Corollary D.2. *There exists a finite set $\mathcal{M}'=\{\lambda_1,\ldots,\lambda_{v_\lambda}\}\subseteq\mathbb{S}_{n_z}(0,1)$ such that for any $\lambda\in\mathbb{R}^{n_z}$ with $\|\lambda\|_2=1$, there exists $\lambda_i\in\mathcal{M}'$ such that $\|\lambda-\lambda_i\|_2\leq 2\epsilon_\lambda$.*

In the following, we consider $\epsilon_\lambda=\sigma_z^2p_z^2/(64b_z^2)=1/a_2$. Notice that $\epsilon_\lambda<1/2$. Accordingly,

$$v_\lambda\leq 544n_z^{2.5}\log(a_2n_z)a_2^{n_z}. \tag{7}$$

Proof. $\epsilon_\lambda\leq 1/64<1/2$ because $p_z\leq 1$ and $\sigma_z\leq b_z$ by the definitions of BMSB and b_z. Then, the bound on v_λ follows from Theorem D.1. □

Lemma D.3. *There exists a finite set $\mathcal{M}=\{\gamma_1,\ldots,\gamma_{v_\gamma}\}\subseteq\mathbb{S}_{n_x\times n_z}(0,1)$ such that for any $\gamma\in\mathbb{R}^{n_x\times n_z}$ and $\|\gamma\|_F=1$, there exists $\gamma_i\in\mathcal{M}$ such that $\|\gamma-\gamma_i\|_F\leq 2\epsilon_\gamma$. Consider $\epsilon_\gamma=\frac{a_1}{4b_z\sqrt{n_x}}=1/a_4$. Notice that $\epsilon_\gamma<1/2$. Accordingly,*

$$v_\gamma\leq 544n_x^{2.5}n_z^{2.5}\log(a_4n_xn_z)a_4^{n_zn_x}.$$

Proof. The proof is basically by mapping the matrices to vectors based on matrix vectorization, then mapping the vectors back to matrices. These two mappings are isomorphism.

Specifically, consider a closed unit ball in $\mathbb{R}^{n_xn_z}$. There exist v_{ϵ,n_xn_z} smaller closed balls to cover it, denoted by $\mathbb{B}_1,\ldots,\mathbb{B}_{v_{\epsilon,n_xn_z}}$. Consider the non-empty sets from $\mathbb{B}_1\cap\mathbb{S}_{n_xn_z}(0,1),\ldots,\mathbb{B}_{v_{\epsilon,n_xn_z}}\cap\mathbb{S}_{n_xn_z}(0,1)$. For any $1\leq i\leq v_{\epsilon,n_xn_z}$, if $\mathbb{B}_i\cap\mathbb{S}_{n_xn_z}(0,1)\neq\emptyset$, select a point $\text{vec}(\gamma)\in\mathbb{B}_i\cap\mathbb{S}_{n_xn_z}(0,1)$. Notice that $\|\text{vec}(\gamma)\|_2=1$. In this way, we construct a finite sequence $\{\text{vec}(\gamma_1),\ldots,\text{vec}(\gamma_{v_\gamma})\}$ where $v_\gamma\leq v_{\epsilon_\gamma,n_xn_z}$.[12]

For any $\gamma\in\mathbb{R}^{n_x\times n_z}$, we have $\text{vec}(\gamma)\in\mathbb{R}^{n_xn_z}$ and $\|\text{vec}(\gamma)\|_2=1$. Hence, there exists $1\leq i\leq v_\gamma$ such that $\text{vec}(\gamma)\in\mathbb{B}_i\cap\mathbb{S}_{n_xn_z}(0,1)$. Hence, $\|\text{vec}(\gamma)-\text{vec}(\gamma_i)\|_2\leq 2\epsilon_\gamma$. Moreover, $\|\gamma_i\|_F=\|\text{vec}(\gamma_i)\|_2=1$. Therefore, $\|\gamma_i-\gamma\|_F\leq 2\epsilon_\gamma$. So the set $\mathcal{M}=\{\gamma_1,\ldots,\gamma_{v_\gamma}\}$ satisfies our requirement. □

D.2. Proof of Lemma 4.1

Essentially, Lemma 4.1 shows that PE holds with high probability under the BMSB condition. This result has been established in Proposition 2.5 in (Simchowitz et al., 2018), though in a different form. The rest of this subsection will prove the PE condition needed in this paper based on Proposition 2.5 in (Simchowitz et al., 2018).

Firstly, we review Proposition 2.5 in (Simchowitz et al., 2018) for the convenience of the reader.

Theorem D.4 (Proposition 2.5 in (Simchowitz et al., 2018) when $k=1$)**.** *Let $\{Z_t\}_{t\geq 1}$ be an $\{\mathcal{F}_t^Z\}_{t\geq 1}$-adapted random process taking values in $\mathbb{R}$. Z_0 is given. If $\{Z_t\}_{t\geq 0}$ is $(1,v,p)$-BMSB, then*

$$\mathbb{P}(\sum_{t=1}^T Z_t^2\leq v^2p^2T/8)\leq\exp(-Tp^2/8)$$

[12] Here, without loss of generality, we consider $\mathbb{B}_1\cap\mathbb{S}_{n_xn_z}(0,1),\ldots,\mathbb{B}_{v_\gamma}\cap\mathbb{S}_{n_xn_z}(0,1)$ are not empty.

Next, we prove the PE in one segment of data sequence.

Lemma D.5 (Probability of PE in one segment)**.** *For any $m \geq 1$, for any $k \geq 0$, we have*

$$\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \succ (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) \geq 1 - v_\lambda \exp(-m p_z^2/8))$$

Proof. Consider $\mathcal{M}' = \{\lambda_1, \ldots, \lambda_{v_\lambda}\}$ defined in Corollary D.2. For any $\lambda_i \in \mathcal{M}'$, $\lambda_i^\top z_t$ satisfies the $(1, \sigma_z, p_z)$-BMSB condition. Therefore, by Theorem D.4, we have

$$\mathbb{P}(\sum_{t=1}^{T} \lambda_i^\top z_t z_t^\top \lambda_i \leq \sigma_z^2 p_z^2 T/8) \leq \exp(-T p_z^2/8).$$

Notice that the horizon length T is arbitrary and the starting stage $t = 1$ can also be different because we consider a time-invariant dynamical system in this paper. Therefore, for any $m \geq 1, k \geq 0$, for any $\lambda_i \in \mathcal{M}'$, we have

$$\mathbb{P}(\sum_{i=1}^{m} \lambda_i^\top z_{km+i} z_{km+i}^\top \lambda_i \leq \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km}) \leq \exp(-m p_z^2/8),$$

where we condition on $\mathcal{F}_{km}$ to make sure z_{km} is known under $\mathcal{F}_{km}$, which is required by Theorem D.4.

For arbitrary λ such that $\|\lambda\|_2 = 1$, there exists $\lambda_i \in \mathcal{M}'$ such that $\|\lambda - \lambda'\|_2 \leq 2\epsilon_\lambda$. Therefore, we can bound $\sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda$ by $\sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i$.

$$\begin{aligned}
\sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda &= \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i + \sum_{t=km+1}^{km+m} (\lambda + \lambda_i)^\top z_t z_t^\top (\lambda - \lambda_i) \\
&\geq \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sum_{t=km+1}^{km+m} \|\lambda + \lambda_i\|_2 \|z_t\|_2^2 \|\lambda_i - \lambda\|_2 \\
&\overset{(a)}{\geq} \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sum_{t=km+1}^{km+m} 4 b_z^2 \epsilon_\lambda \\
&= \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - 4 b_z^2 \epsilon_\lambda m \overset{(b)}{\geq} \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i - \sigma_z^2 p_z^2 m/16
\end{aligned}$$

where (a) is by Assumption 2.5, $\|\lambda - \lambda_i\|_2 \leq 2\epsilon_\lambda$, and $\|\lambda\|_2 = \|\lambda_i\|_2 = 1$; and (b) is by choosing $\epsilon_\lambda \leq \sigma_z^2 p_z^2/(64 b_z^2)$.

Therefore, by the definition of positive definiteness and the inequalities above, we can complete the proof by the following.

$$\begin{aligned}
\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \succ (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) &= \mathbb{P}(\forall \|\lambda\|_2 = 1, \sum_{t=km+1}^{km+m} \lambda^\top z_t z_t^\top \lambda > \sigma_z^2 p_z^2 m/16 \mid \mathcal{F}_{km}) \\
&\geq \mathbb{P}(\forall 1 \leq i \leq v_\lambda, \sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i > \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km}) \\
&\geq 1 - \sum_{i=1}^{v_\lambda} \mathbb{P}(\sum_{t=km+1}^{km+m} \lambda_i^\top z_t z_t^\top \lambda_i \leq \sigma_z^2 p_z^2 m/8 \mid \mathcal{F}_{km}) \\
&\geq 1 - v_\lambda \exp(-m p_z^2/8)),
\end{aligned}$$

which completes the proof. □

Now, we are ready for the proof of Lemma 4.1.

Proof of Lemma 4.1. Recall that $\mathcal{E}_2 = \{\frac{1}{m}\sum_{s=1}^m z_{km+s}z_{km+s}^\top \succeq a_1^2 I_{n_z},\ \forall\, 0\leq k\leq \lceil T/m\rceil -1\}$, where $a_1=\sigma_z p_z/4$. Hence

$$\mathcal{E}_2 = \bigcap_{k=0}^{T/m-1}\{\sum_{t=km+1}^{km+m} z_t z_t^\top \succ (\sigma_z^2 p_z^2 m/16) I_{n_z}\}.$$

Therefore,

$$\begin{aligned}\mathbb{P}(\mathcal{E}_2) &\geq 1-\sum_{k=0}^{T/m-1}\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \preceq (\sigma_z^2 p_z^2 m/16) I_{n_z})\\ &\geq 1-\frac{T}{m}v_\lambda \exp(-m p_z^2/8)\\ &= 1-\frac{T}{m}(544 n_z^{2.5}\log(a_2 n_z) a_2^{n_z})\exp(-m p_z^2/8),\end{aligned}$$

where we use Lemma D.5 and the fact that if $\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \preceq (\sigma_z^2 p_z^2 m/16) I_{n_z} \mid \mathcal{F}_{km}) \leq v_\lambda \exp(-m p_z^2/8))$, then $\mathbb{P}(\sum_{t=km+1}^{km+m} z_t z_t^\top \preceq (\sigma_z^2 p_z^2 m/16) I_{n_z}) \leq v_\lambda \exp(-m p_z^2/8))$. □

D.3. Proof of Lemma 4.2

This proof takes four major steps:

(i) Define $b_{i,t}, j_{i,t}, L_{i,k}$.

(ii) Provide a formal definition of $\mathcal{E}_{1,k}$ based on $b_{i,t}, j_{i,t}, L_{i,k}$ and prove a formal version of Lemma 4.3.

(iii) Prove Lemma 4.4.

(iv) Prove Lemma 4.2 by the formal version of Lemma 4.3 and Lemma 4.4.

It is worth mentioning that the formal definition of $\mathcal{E}_{1,k}$ is slightly different from the definition in Lemma 4.3, but we still have $\mathbb{P}(\mathcal{E}_1\cap\mathcal{E}_2)\leq \sum_{i=1}^{v_\gamma}\mathbb{P}(\mathcal{E}_{1,k}\cap\mathcal{E}_2)$, which is the key property that will be used in the proof of Lemma 4.2.

D.3.1. Step (i): Definitions of $b_{i,t}, j_{i,t}, L_{i,k}$.

Recall the discretization of $\mathbb{S}_{n_x\times n_z}(0,1)$ in Lemma D.3, which generates the set $\mathcal{M}=\{\gamma_1,\dots,\gamma_{v_\gamma}\}$. We are going to define $b_{i,t}, j_{i,t}, L_{i,k}$ for $\gamma_i\in\mathcal{M}$ for each $1\leq i\leq v_\gamma$. Notice that $\mathcal{M}$ is a deterministic set of matrices.

Lemma D.6 (Definition of $b_{i,t}, j_{i,t}$). *For any $\gamma_i\in\mathcal{M}$, any $0\leq t\leq T$, there exist $b_{i,t}\in\{-1,1\}$ and $1\leq j_{i,t}\leq n_x$ such that $b_{i,t}, j_{i,t}\in\mathcal{F}(z_t)\subseteq\mathcal{F}_t$ and*

$$\|\gamma_i z_t\|_\infty = b_{i,t}(\gamma_i z_t)^{j_{i,t}}.$$

Note that one way to determine $b_{i,t}, j_{i,t}$ from z_t is by the following: first pick the smallest j such that $|(\gamma_i z_t)^j| = \|\gamma_i z_t\|_\infty$, then let $b_{i,t}=\mathrm{sgn}((\gamma_i z_t)^j)$, where $\mathrm{sgn}(\cdot)$ denotes the sign of a scalar argument.

Proof. For any $\gamma_i\in\mathcal{M}$, any $0\leq t\leq T$, we have

$$\|\gamma_i z_t\|_\infty = \max_{1\leq j\leq n_x}\max_{b\in\{-1,1\}} b(\gamma_i z_t)^j$$

Hence, there exist $b_{i,t}, j_{i,t}$ such that $\|\gamma_i z_t\|_\infty = b_{i,t}(\gamma_i z_t)^{j_{i,t}}$. Further, $b_{i,t}, j_{i,t}$ only depend on γ_i and z_t, so they are $\mathcal{F}(z_t)$-measurable, and $\mathcal{F}(z_t)\subseteq\mathcal{F}_t$. □

Lemma D.7 (Definition of stopping times $L_{i,k}$). *Let $\eta=\frac{a_1}{\sqrt{n_x}}$. For any $\gamma_i\in\mathcal{M}$, any $0\leq k\leq T/m-1$, we can define a random time index $1\leq L_{i,k}\leq m+1$ by*

$$L_{i,k}=\min(m+1, \min\{l\geq 1: \|\gamma_i z_{km+l}\|_\infty\geq\eta\}).$$

Then, we have $1\leq L_{i,k}\leq m+1$. Further, for any $1\leq l\leq m$, $\{L_{i,k}=l\}\in\mathcal{F}_{km+l}$, and $\{L_{i,k}=m+1\}\in\mathcal{F}_{km+m}\subseteq\mathcal{F}_{km+m+1}$. In other words, $L_{i,k}$ is a stopping time with respect to filtration $\{F_{km+l}\}_{l\geq 1}$.

Proof. For any i and any k, it is straightforward to see that $L_{i,k}$ is well-defined and $1 \leq L_{i,k} \leq m+1$.

When $L_{i,k} = l \leq m$, this is equivalent with $\|\gamma_i z_{km+l}\|_\infty \geq \eta$ but $\|\gamma_i z_{km+s}\| < \eta$ for $1 \leq s < l$. Notice that this event is only determined by $z_{km+l}, \ldots, z_{km+1}$, so $\{L_{i,k} = l\} \in \mathcal{F}_{km+l}$.

When $L_{i,k} = m+1$, this is equivalent with $\|\gamma_i z_{km+s}\| < \eta$ for $1 \leq s \leq m$. Notice that this event is only determined by $z_{km+m}, \ldots, z_{km+1}$, so $\{L_{i,k} = m+1\} \in \mathcal{F}_{km+m}$.

Therefore, by definition, $L_{i,k}$ is a stopping time with respect to filtration $\{\mathcal{F}_{km+l}\}_{l \geq 1}$. □

D.3.2. Step (ii): a formal version of Lemma 4.3 and its proof

Lemma D.8 (Discretization of $\mathcal{E}_1 \cap \mathcal{E}_2$ (Formal version of Lemma 4.3))**.** *Let $\mathcal{M} = \{\gamma_1, \ldots, \gamma_{v_\gamma}\}$ be an ϵ_γ-net of $\{\gamma : \|\gamma\|_F = 1\}$ as defined in Lemma D.3, where $\epsilon_\gamma = \min(\frac{a_1}{4b_z\sqrt{n_x}}, 1)$, $v_\gamma = \tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}$, and $a_4 = \frac{4b_z\sqrt{n_x}}{a_1}$. Define*

$$\mathcal{E}_{1,i} = \{\exists\, \gamma \in \Gamma_T,\ s.t.\ b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1 \delta}{4\sqrt{n_x}},\ \forall\, k \geq 0\}.$$

Then, we have

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2).$$

The rest of this subsubsection is dedicated to the proof of Lemma D.8. As an overview: firstly, we will discuss the implications of $\mathcal{E}_2$ on $\gamma_i \in \mathcal{M}$. Then, we discuss the implications of $\mathcal{E}_2$ on any γ. Lastly, we prove Lemma D.8 by combining the implications of $\mathcal{E}_2$ on any γ and $\|\gamma\|_F \geq \delta/2$.

Lemma D.9 (The implication of $\mathcal{E}_2$ on γ_i)**.** *If $\mathcal{E}_2$ happens, then for any $\gamma_i \in \mathcal{M}$, any $0 \leq k \leq T/m - 1$, we have*

$$\max_{1 \leq s \leq m} \|\gamma_i z_{km+s}\|_\infty \geq \frac{a_1}{\sqrt{n_x}}.$$

Therefore, almost surely, we have $1 \leq L_{i,k} \leq m$ and

$$b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}.$$

Proof. If $\mathcal{E}_2$ happens, then by definition, we have

$$\frac{1}{m}\sum_{s=1}^{m} z_{km+s} z_{km+s}^\top \succeq a_1^2 I_{n_z},$$

for all $0 \leq k \leq T/m - 1$.

Now, for any $\gamma_i \in \mathcal{M}$, we have that

$$\frac{1}{m}\sum_{s=1}^{m} \gamma_i z_{km+s} z_{km+s}^\top \gamma_i^\top \succeq a_1^2 \gamma_i \gamma_i^\top. \tag{8}$$

Therefore, by taking trace at each side of (8), we obtain

$$\frac{1}{m}\sum_{s=1}^{m} \mathrm{tr}(\gamma_i z_{km+s} z_{km+s}^\top \gamma_i^\top) \geq a_1^2 \mathrm{tr}(\gamma_i \gamma_i^\top) \tag{9}$$

Since $\gamma_i \in \mathbb{S}_{n_x \times n_z}(0,1)$, we have $\|\gamma_i\|_F = 1$, so $\mathrm{tr}(\gamma_i\gamma_i^\top) = \mathrm{tr}(\gamma_i^\top \gamma_i) = \|\gamma_i\|_F^2 = 1$. Further, we have

$$\mathrm{tr}(\gamma_i z_{km+s} z_{km+s}^\top \gamma_i^\top) = \mathrm{tr}(z_{km+s}^\top \gamma_i^\top \gamma_i z_{km+s}) = z_{km+s}^\top \gamma_i^\top \gamma_i z_{km+s} = \|\gamma_i z_{km+s}\|_2^2.$$

Consequently, we have

$$\frac{1}{m}\sum_{s=1}^{m}\|\gamma_i z_{km+s}\|_2^2 \geq a_1^2$$

for all k.

By the pigeonhole principle, we have that

$$\max_{1\leq s\leq m}\|\gamma_i z_{km+s}\|_2^2 \geq a_1^2.$$

This is equivalent with $\max_{1\leq s\leq m}\|\gamma_i z_{km+s}\|_2 \geq a_1$.

Notice that $\|\gamma_i z_{km+s}\|_2 \leq \sqrt{n_x}\|\gamma_i z_{km+s}\|_\infty$, so $\max_{1\leq s\leq m}\sqrt{n_x}\|\gamma_i z_{km+s}\|_\infty \geq a_1$, which completes the proof of the first inequality in the lemma statement.

Next, we prove the second inequality in the lemma statement. Notice that by the definition of $L_{i,k}$ in Lemma D.7 and by $\eta = \frac{a_1}{\sqrt{n_x}}$, we have $1 \leq L_{i,k} \leq m$ and $\|\gamma_i z_{km+L_{i,k}}\|_\infty \geq \frac{a_1}{\sqrt{n_x}}$ for all k. Further, by Lemma D.6, we have $\|\gamma_i z_{km+L_{i,k}}\|_\infty = b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}}$ almost surely. Hence, we have $b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}$, which completes the proof.

□

Lemma D.10 (The implication of $\mathcal{E}_2$ on γz_t)**.** *If $\mathcal{E}_2$ happens, then for any $\gamma \in \mathbb{R}^{n_x\times n_z}$, there exists $1\leq i\leq v_\gamma$, such that*

$$b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}}\|\gamma\|_F,$$

for all $0\leq k\leq T/m-1$.

Proof. Firstly, when $\gamma = 0$, the inequality holds because both sides are 0.

Next, when $\gamma\neq 0$, it suffices to prove $b_{i,km+L_{i,k}}(\frac{\gamma}{\|\gamma\|_F}z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}}$. Therefore, we will only consider $\gamma\in\mathbb{S}_{n_x\times n_z}(0,1)$. By Lemma D.3, there exists $\gamma_i\in\mathcal{M}$ such that $\|\gamma-\gamma_i\|_F\leq 2\epsilon_\gamma = \min(\frac{a_1}{2b_z\sqrt{n_x}},2)$. Notice that by Lemma D.9, if $\mathcal{E}_2$ happens, for all k, we have

$$b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{\sqrt{n_x}}.$$

Therefore,

$$\begin{aligned}
b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} &= b_{i,km+L_{i,k}}(\gamma_i z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \\
&\quad - b_{i,km+L_{i,k}}((\gamma_i-\gamma)z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \\
&\geq \frac{a_1}{\sqrt{n_x}} - |b_{i,km+L_{i,k}}((\gamma_i-\gamma)z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}}| \\
&\geq \frac{a_1}{\sqrt{n_x}} - \|(\gamma_i-\gamma)z_{km+L_{i,k}}\|_2 \\
&\geq \frac{a_1}{\sqrt{n_x}} - \|\gamma_i-\gamma\|_2\|z_{km+L_{i,k}}\|_2 \\
&\geq \frac{a_1}{\sqrt{n_x}} - 2\epsilon_\gamma b_z \geq \frac{a_1}{2\sqrt{n_x}}
\end{aligned}$$

□

Proof of Lemma D.8. By Lemma D.10, under $\mathcal{E}_2$, for any $\gamma\in\mathbb{R}^{n_x\times n_z}$, there exists $1\leq i\leq v_\gamma$, such that

$$b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}}\|\gamma\|_F,$$

for all $0 \leq k \leq T/m - 1$. Therefore, if $\mathcal{E}_1 \cap \mathcal{E}_2$ happens, there exists $\gamma \in \Gamma_T$ and a corresponding i, such that

$$b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1}{2\sqrt{n_x}}\|\gamma\|_F \geq \frac{a_1\delta}{4\sqrt{n_x}}.$$

Therefore,

$$\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) \leq \mathbb{P}(\bigcup_{i=1}^{v_\gamma} \mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2),$$

which completes the proof. □

D.3.3. Proof of Lemma 4.4

Notice that Lemma 4.4 states two inequalities: in the following, we will first prove the first inequality $\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}(\cap_{k=0}^{T/m-1} G_{i,k})$, then prove the second inequality on $\mathbb{P}(G_{i,k} \mid \cap_{k'=0}^{k-1} G_{i,k'})$.

Lemma D.11 (Bound $\mathcal{E}_{1,i} \cap \mathcal{E}_2$ by $G_{i,k}$)**.** *Under the conditions in Lemma 4.4, for any $1 \leq i \leq v_\gamma$, we have*

$$\mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \leq \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}).$$

Proof. Firstly, for any $\gamma \in \Gamma_T$, we have $\|w_t - \gamma z_t\|_\infty \leq w_{\max}$ for all $t \geq 0$. This suggests that, for any $1 \leq j \leq n_x$, we have

$$-w_{\max} \leq w_t^j - (\gamma z_t)^j \leq w_{\max}.$$

Hence, we have $b(\gamma z_t)^j \leq b w_t^j + w_{\max}$ for any $b \in \{-1, 1\}$, $1 \leq j \leq n_x$, and $t \geq 0$.

Next, by $\mathcal{E}_{1,i}$, there exists $\gamma \in \Gamma_T$ such that $b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1\delta}{4\sqrt{n_x}}$ for all $k \geq 0$. Therefore, $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}$ for all k.

Finally, $\mathcal{E}_{1,i} \cap \mathcal{E}_2$ implies that $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}$ and $\frac{1}{m}\sum_{s=1}^m z_{km+s} z_{km+s}^\top \succeq a_1^2 I_{n_z}$ for all k, which is $\bigcap_k G_{i,k}$ by the definition of $G_{i,k}$.

□

Lemma D.12 (Bound on $\mathbb{P}(G_{i,k} \mid \cap_{k'=0}^{k-1} G_{i,k'})$)**.** *Under the conditions in Lemma 4.4, for any $1 \leq i \leq v_\gamma$ and any $k \geq 0$, we have*

$$\mathbb{P}(G_{i,k} \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \leq 1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}).$$

Proof. Firstly, notice that when $\frac{1}{m}\sum_{s=1}^m z_{km+s} z_{km+s}^\top \succeq a_1^2 I_{n_z}$, we have $1 \leq L_{i,k} \leq m$ by the proof of Lemma D.9. Therefore, we have

$$\begin{aligned}
\mathbb{P}(G_{i,k} \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) &\leq \mathbb{P}(b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}, 1 \leq L_{i,k} \leq m \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \\
&\leq \sum_{l=1}^m \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}, L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \\
&\leq \sum_{l=1}^m \mathbb{P}(b_{i,km+l} w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \mathbb{P}(L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'}) \\
&\overset{(a)}{\leq} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}})) \sum_{l=1}^m \mathbb{P}(L_{i,k} = l \mid \bigcap_{k'=0}^{k-1} G_{i,k'})
\end{aligned}$$

$$\leq 1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}})$$

The inequality (c) is proved in the following:

$$\begin{aligned}
&\mathbb{P}(b_{i,km+l}w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \\
&= \int_{v_{0:km+l}} \mathbb{P}(b_{i,km+l}w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}}, w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \mathrm{d}v_{0:km+l} \\
&= \int_{v_{0:km+l}\in S_{km+l}} \mathbb{P}(b_{i,km+l}w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}} \mid w_{0:km+l} = v_{0:km+l}) \\
&\qquad\qquad \times \mathbb{P}(w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \mathrm{d}v_{0:km+l} \\
&\overset{(b)}{\leq} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}})) \int_{v_{0:km+l}\in S_{km+l}} \mathbb{P}(w_{0:km+l} = v_{0:km+l} \mid L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}) \mathrm{d}v_{0:km+l} \\
&= 1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}),
\end{aligned}$$

where we define a shorthand notation $w_{0:km+l} = (w_0, \ldots, w_{km+l-1})$, and we use $v_{0:km+l}$ to denote a realization of $w_{0:km+l}$, then we define the set of values of $w_{0:km+l}$ as S_{km+l} such that $L_{i,k} = l, \bigcap_{k'=0}^{k-1} G_{i,k'}$ holds. Notice that $L_{i,k} = l$ can be determined by a set of values of $w_{0:km+l}$ because $L_{i,k}$ is a stopping time of $\{F_{km+l}\}_{l\geq 1}$ and thus $\{L_{i,k} = l\} \in \mathcal{F}_{km+l}$. The inequality (b) above is because of the following: firstly, notice that $b_{i,km+l}, j_{i,km+l} \in \mathcal{F}_{km+l}$, so $b_{i,km+l}, j_{i,km+l}$ are deterministic values when $w_{0:km+l} = v_{0:km+l}$. Further, since w_{km+l} is independent of $w_{0:km+l}$, we have $\mathbb{P}(w_{max} + bw_{km+l}^j \geq \epsilon \mid w_{0:km+l} = v_{0:km+l}) \leq 1 - q_w(\epsilon)$ for any deterministic b, j and any $\epsilon > 0$ by Assumption 2.6. Hence, we have $\mathbb{P}(b_{i,km+l}w_{km+l}^{j_{i,km+l}} + w_{\max} \geq \frac{a_1\delta}{4\sqrt{n_x}} \mid w_{0:km+l} = v_{0:km+l}) \leq 1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}})$. □

D.3.4. Proof of Lemma 4.2

The proof is by leveraging Lemma D.8 and Lemma 4.4.

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) &\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{1,i} \cap \mathcal{E}_2) \\
&\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}) \\
&= \sum_{i=1}^{v_\gamma} \mathbb{P}(G_{i,0})\mathbb{P}(G_{i,1} \mid G_{i,0}) \cdots \mathbb{P}(G_{i,T/m-1} \mid \bigcap_{k=0}^{T/m-2} G_{i,k}) \\
&\leq \sum_{i=1}^{v_\gamma} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{T/m} \\
&\leq 544 n_x^{2.5} n_z^{2.5} \log(a_4 n_x n_z) a_4^{n_z n_x} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{T/m}.
\end{aligned}$$

D.4. A more precise upper bound for Theorem 3.1

By the proof of Lemma 4.1 and Lemma 4.2 above, we have

$$\mathbb{P}(\mathrm{diam}(\Theta_T) > \delta) \leq 544 \frac{T}{m} n_z^{2.5} \log(a_2 n_z) a_2^{n_z} \exp(-a_3 m) + 544 n_x^{2.5} n_z^{2.5} \log(a_4 n_x n_z) a_4^{n_z n_x} (1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{T/m} \quad (10)$$

E. Proof of Corollary 3.2

The proof involves two parts. Firstly, we will show that Term 1 $\leq \epsilon$ under our choice of m. Secondly, we will let Term 2 $= \epsilon$, then we will show $\delta \leq \tilde{O}(n_x^{1.5} n_z^2/T)$, which completes the proof.

Step 1: show Term 1 $\leq \epsilon$. Notice that when $m \geq \frac{1}{a_3}(\log(\frac{T}{\epsilon}) + n_z \log(a_2) + 2.5\log(n_z) + \log\log(a_2 n_z) + 7) = O(n_z + \log T + \log(1/\epsilon))$, we have $T\tilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m) \leq \epsilon$. Since $m \geq 1$, we obtain Term 1 $\leq \epsilon$.

Step 2: let Term 2 $= \epsilon$ and show $\delta \leq \tilde{O}(n_x^{1.5} n_z^2/T)$. Let Term 2 $= \epsilon$, then we have $(1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{T/m} = \frac{\epsilon}{\tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}$. Then, we obtain $(1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}})) = \left(\frac{\epsilon}{\tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}$, which is equivalent with

$$q_w(\frac{a_1\delta}{4\sqrt{n_x}}) = 1 - \left(\frac{\epsilon}{\tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}.$$

When $q_w(\frac{a_1\delta}{4\sqrt{n_x}}) = O(\frac{a_1\delta}{4\sqrt{n_x}})$, we obtain

$$\begin{aligned}
\delta &= O(\frac{4\sqrt{n_x}}{a_1})\left(1 - \left(\frac{\epsilon}{\tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}\right) \\
&\leq O(\frac{-4\sqrt{n_x}}{a_1})\log\left(\left(\frac{\epsilon}{\tilde{O}(n_x^{2.5} n_z^{2.5}) a_4^{n_x n_z}}\right)^{m/T}\right) \\
&= O(\frac{4\sqrt{n_x} m}{a_1 T})(\log(1/\epsilon) + n_x n_z + \log(n_x n_z)) \\
&= \tilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right).
\end{aligned}$$

Step 3: prove Corollary 3.2. By leveraging the bounds above and Theorem 3.1, we have $\mathbb{P}(\text{diam}(\Theta_T) \leq \tilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right)) \geq \mathbb{P}(\text{diam}(\Theta_T) \leq \delta) \geq 1 - 2\epsilon$.

Since $\theta^* \in \Theta_T$ by definition, for any $\hat{\theta}_T \in \Theta_T$, we have $\|\hat{\theta}_T - \theta^*\|_F \leq \text{diam}(\Theta_T) \leq \tilde{O}\left(\frac{n_x^{1.5} n_z^2}{T}\right)$ with probability at least $1 - 2\epsilon$.

F. Proof of Corollary 3.3

We provide a formal version of Corollary 3.3 and its proof below.

Corollary F.1 (Convergence rate when $B^* = 0$ (formal version))**.** *When A^* is (κ, ρ)-stable, i.e., $\|(A^*)^t\|_2 \leq \kappa(1-\rho)^t$ for all t with $\rho < 1$, for any $m > 0$ and any $\delta > 0$, when $T > m$, we have*

$$\mathbb{P}(\text{diam}(\mathbb{A}_T) > \delta) \leq \frac{T}{m}\tilde{O}(n_x^{2.5})a_2^{n_x}\exp(-a_3 m) + \tilde{O}(n_x^5)a_4^{n_x^2}(1 - q_w(\frac{a_1\delta}{4\sqrt{n_x}}))^{\lceil T/m \rceil}$$

where $b_x = \kappa\|x_0\|_2 + \kappa\sqrt{n_x}/\rho$, $p_x = 1/192$, $\sigma_x = \sqrt{\lambda_{\min}(\Sigma_w)/2}$, $a_1 = \frac{\sigma_x p_x}{4}$, $a_2 = \frac{64 w_{\max}}{\sigma_x^2 p_x^2}$, $a_3 = \frac{p_x^2}{8}$, $a_4 = \frac{4 b_x \sqrt{n_x}}{a_1}$.

Consequently, when the distribution of w_t satisfies $q_w(\epsilon) = O(\epsilon)$, e.g. uniform or truncated Gaussian, we have $\|\hat{\theta} - \theta_*\| \leq \tilde{O}(n_x^{3.5}/T)$.

The proof of Corollary 3.3 is exactly the same as the proofs of Theorem 3.1 and Corollary 3.2. When A^* is stable, we can show that $\|x_t\|_2 \leq b_x$ for all t. Further, by (Dean et al., 2019b), the sequence $\{x_t\}_{t\geq 0}$ satisfies the $(1, \sigma_x, p_x)$-BMSB condition. Therefore, we complete the proof.

G. Proof of Theorem 3.4

Specifically, we define $\epsilon_0 = \frac{4\sqrt{n_x}}{a_1}(\hat{w}_{\max} - w_{\max})$.

The proof is similar to the proof of Theorem 3.1. Firstly, we define $\hat{\Gamma}_T$ as a translation of the set $\hat{\Theta}_T$:

$$\hat{\Gamma}_t = \bigcap_{s=0}^{t-1} \{\gamma : \|w_s - \gamma z_s\|_\infty \leq \hat{w}_{\max}\}, \quad \forall\, t \geq 0. \tag{11}$$

Notice that

$$\hat{\Theta}_T = \theta^* + \hat{\Gamma}_T$$

by considering $\gamma = \hat{\theta} - \theta^*$. Therefore, we can upper bound our goal event $\{\text{diam}(\hat{\Theta}_T) > \delta + \epsilon_0\}$ by the event $\mathcal{E}_3$ defined below.

$$\mathbb{P}(\text{diam}(\hat{\Theta}_T) > \delta + \epsilon_0) \leq \mathbb{P}(\mathcal{E}_3), \text{ where } \mathcal{E}_3 := \{\exists\, \gamma \in \hat{\Gamma}_T, \text{ s.t. } \|\gamma\|_F \geq \frac{\delta + \epsilon_0}{2}\}. \tag{12}$$

Next, notice that

$$\mathbb{P}(\text{diam}(\hat{\Theta}_T) > \delta + \epsilon_0) \leq \mathbb{P}(\mathcal{E}_3) \leq \mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) + \mathbb{P}(\mathcal{E}_2^c)$$

By Lemma 4.1, we have already shown $\mathbb{P}(\mathcal{E}_2^c) \leq$ Term 1. So we only need to discuss $\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2)$.

Lemma G.1.

$$\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) \leq \text{Term 2}$$

Proof. Firstly, define

$$\mathcal{E}_{3,i} = \{\exists\, \gamma \in \hat{\Gamma}_T, \text{ s.t. } b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1(\delta + \epsilon_0)}{4\sqrt{n_x}}, \forall\, k \geq 0\}.$$

We have $\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{3,i} \cap \mathcal{E}_2)$ based on the same proof ideas of Lemma D.8.

Next, we will show that

$$\Pr(\mathcal{E}_{3,k} \cap \mathcal{E}_2) \leq \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}) \tag{13}$$

This is because for any $\gamma \in \hat{\Gamma}_T$, we have $b(\gamma z_t)^j \leq b w_t^j + \hat{w}_{\max}$ for any $b \in \{-1, 1\}$, $1 \leq j \leq n_x$, and $t \geq 0$. By $\mathcal{E}_{3,i}$, there exists $\gamma \in \hat{\Gamma}_T$ such that $b_{i,km+L_{i,k}}(\gamma z_{km+L_{i,k}})^{j_{i,km+L_{i,k}}} \geq \frac{a_1(\delta+\epsilon_0)}{4\sqrt{n_x}}$ for all $k \geq 0$. Thus, $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + \hat{w}_{\max} \geq \frac{a_1(\delta+\epsilon_0)}{4\sqrt{n_x}}$ for all k. Notice that this is equivalent with $b_{i,km+L_{i,k}} w_{km+L_{i,k}}^{j_{i,km+L_{i,k}}} + w_{\max} \geq \frac{a_1 \delta}{4\sqrt{n_x}}$ for all k because $\epsilon_0 = \frac{4\sqrt{n_x}}{a_1}(\hat{w}_{\max} - w_{\max})$. In this way, we can prove (13).

Finally, we can complete the proof by the following.

$$\begin{aligned}
\mathbb{P}(\mathcal{E}_3 \cap \mathcal{E}_2) &\leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\mathcal{E}_{3,i} \cap \mathcal{E}_2) \leq \sum_{i=1}^{v_\gamma} \mathbb{P}(\bigcap_{k=0}^{T/m-1} G_{i,k}) \\
&= \sum_{i=1}^{v_\gamma} \mathbb{P}(G_{i,0}) \mathbb{P}(G_{i,1} \mid G_{i,0}) \cdots \mathbb{P}(G_{i,T/m-1} \mid \bigcap_{k=0}^{T/m-2} G_{i,k}) \\
&\leq \sum_{i=1}^{v_\gamma} (1 - q_w(\frac{a_1 \delta}{4\sqrt{n_x}}))^{T/m} \leq \text{Term 1}
\end{aligned}$$

where the second last inequality is by Lemma D.12 and the last inequality uses the definition of v_γ in Lemma D.3. □

H. Proofs of Theorem 3.5, Corollary 3.6, and Theorem 3.7

This section provides proofs of the main results related to the SME with unknown $w_{\max}$ as discussed in Section 3.2. Namely, Theorem 3.5 and Corollary 3.6 provide the rate of convergence of the estimator $\bar{w}_{\max}^{(T)}$ defined in (4) to $w_{\max}$, and Theorem 3.7 states the rate of convergence of UCB-SME algorithm introduced in (5).

For ease of notation, we introduce the following function indexed by the time horizon $T > 0$,

$$W_T : \theta \mapsto \max_{0 \le t \le T-1} \|x_{t+1} - \theta z_t\|_\infty. \tag{14}$$

The estimator $\bar{w}_{\max}^{(T)}$ is simply the infimum of this function, i.e., $\bar{w}_{\max}^{(T)} = \inf_\theta W_T(\theta)$.

H.1. Proof of Theorem 3.5

The proof of Theorem 3.5 involves two steps:

- *Step 1:* We demonstrate that the learning error of $w_{\max}$ incurred by the estimator $\bar{w}_{\max}^{(T)}$ is governed by the diameter of the uncertainty set Θ_T and the minimum learning error achievable if θ^* were known.
- *Step 2:* We then provide an upper bound the probability of learning error exceeding a fixed threshold.

Before we proceed with the proof of Theorem 3.5, we present the the following technical lemma.

Lemma H.1. *Consider the sequence of functions $\{W_T\}_{T>0}$ defined in (14). The following holds:*

i. W_T is convex in $\mathbb{R}^{n_x \times n_z}$,

ii. The sequence $\{\inf_\theta W_T(\theta)\}_{T>0}$ is bounded and monotonically non-decreasing, i.e.,

$$0 \le \inf_\theta W_T(\theta) \le \inf_\theta W_{T+1}(\theta) \le w_{\max},$$

for all $T > 0$,

iii. W_T attains its minimum in Θ_T, i.e., $\arg\min_\theta W_T(\theta) \subset \Theta_T$.

Proof. $(i.)$ For $0 \le t \le T-1$, the function $\theta \mapsto \|x_{t+1} - \theta z_t\|_\infty$ is convex due to convexity of norms. Since the maximum of convex functions is convex (Boyd & Vandenberghe, 2004), convexity of W_T follows.
$(ii.)$ Notice that W_{T+1} can be defined in terms of W_T recursively as $W_{T+1}(\theta) = \max\left(W_T(\theta),\ \|x_{T+1} - \theta z_T\|_\infty\right)$. Thus, $W_T(\theta) \le W_{T+1}(\theta)$ for all $\theta \in \mathbb{R}^{n_x \times n_z}$, implying monotonicity of $\{\inf_\theta W_T(\theta)\}_{T>0}$. To see boundedness, first notice that

$$W_T(\theta^*) = \max_{0 \le t \le T-1} \|x_{t+1} - \theta^* z_t\|_\infty = \max_{0 \le t \le T-1} \|w_t\|_\infty \le w_{\max},$$

since $x_{t+1} = \theta^* z_t + w_t$. Therefore, for any $T > 0$, we have that

$$\inf_\theta W_T(\theta) = \inf_\theta \max_{0 \le t \le T-1} \|x_{t+1} - \theta z_t\|_\infty \le \max_{0 \le t \le T-1} \|x_{t+1} - \theta^* z_t\|_\infty \le w_{\max}$$

.
$(iii.)$ First, we show that W_T attains its minimum on $\mathbb{R}^{n_x \times n_z}$. If $z_t = 0$ for $t \in [T]$, then W_T is a constant function and any $\theta \in \mathbb{R}^{n_x \times n_z}$ is a minimum of W_T. Now, suppose $z_t \ne 0$ for some $t \in [T]$. Then, W_T diverges at the infinity, i.e., $\lim_{k \to \infty} W_T(\theta_k) = \infty$ for any sequence $\{\theta_k\}_{k \in \mathbb{N}}$ such that $\|\theta_k\| \to \infty$ as $k \to \infty$. Since W_T is convex and bounded below with finite infimum, there exists a global minimizer $\bar{\theta}_T \in \mathbb{R}^{n_x \times n_z}$ such that $W_T(\bar{\theta}_T) = \inf_\theta W_T(\theta) = \bar{w}_{\max}^{(T)}$. Furthermore, by (ii), we have that $\|x_{t+1} - \bar{\theta}_T z_t\|_\infty \le w_{\max}$ for all $t \in [T]$ and any global minimizer $\bar{\theta}_T \in \arg\min_\theta W_T(\theta)$, hence $\bar{\theta}_T \in \Theta_T$ by definition. □

Step 1 of the Proof of Theorem 3.5: We first show that the error margin of the estimate $\bar{w}_{\max}^{(T)}$ from $w_{\max}$ is governed by the sum of two factors: (i) the diameter of Θ_T, which arises due to the lack of knowledge of θ^*, and (ii) the minimum learning error achievable if θ^* were known, namely

$$0 \leq w_{\max} - \bar{w}_{\max}^{(T)} \leq b_z \text{diam}(\Theta_T) + w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty. \tag{15}$$

First, $0 \leq w_{\max} - \bar{w}_{\max}^{(T)}$ is simply due to Lemma H.1. Next, we prove the second inequality $w_{\max} - \bar{w}_{\max}^{(T)} \leq b_z \text{diam}(\Theta_T) + w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty$. By Lemma H.1, there exists $\bar{\theta}_T \in \Theta_T$ such that $W_T(\bar{\theta}_T) = w_{\max}$ and

$$\begin{aligned}
w_{\max} &= \max_{0 \leq t \leq T-1} \|x_{t+1} - \bar{\theta}_T z_t\|_\infty, \\
&= \max_{0 \leq t \leq T-1} \|x_{t+1} - \theta^* z_t + (\theta^* - \bar{\theta}_T) z_t\|_\infty, \\
&\geq \max_{0 \leq t \leq T-1} \left(\|x_{t+1} - \theta^* z_t\|_\infty - \|(\theta^* - \theta_k) z_t\|_\infty \right),
\end{aligned}$$

where the inequality is due to reverse triangle inequality. Furthermore, by using the equivalence of ℓ_2 and ℓ_∞ norms, i.e., $\|x\|_2 \leq \|x\|_\infty$ for $x \in \mathbb{R}^{n_x}$, we bound $w_{\max}$ further below by

$$\begin{aligned}
w_{\max} &\geq \max_{0 \leq t \leq T-1} \left(\|x_{t+1} - \theta^* z_t\|_\infty - \|(\theta^* - \bar{\theta}_T) z_t\|_2 \right), \\
&\geq \max_{0 \leq t \leq T-1} \left(\|x_{t+1} - \theta^* z_t\|_\infty - \|\theta^* - \bar{\theta}_T\|_2 \|z_t\|_2 \right), \\
&\geq \max_{0 \leq t \leq T-1} \|w_t\|_\infty - b_z \text{diam}(\Theta_T),
\end{aligned}$$

where the second inequality is due to $\|\theta^* - \bar{\theta}_T\|_2 := \sup_{z \neq 0} \frac{\|(\theta^* - \bar{\theta}_T) z\|_2}{\|z\|_2} \leq \frac{\|(\theta^* - \bar{\theta}_T) z_t\|_2}{\|z_t\|_2}$ and the third inequality follows from the assumption $\|z_t\|_2 \leq b_z$, the equivalence of Frobenius and spectral norms $\|\theta^* - \bar{\theta}_T\|_2 \leq \|\theta^* - \bar{\theta}_T\|_F$, and $\theta^*, \bar{\theta}_T \in \Theta_T$. Consequently,

$$w_{\max} - \bar{w}_{\max}^{(T)} \leq w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty + b_z \text{diam}(\Theta_T).$$

This completes the proof of the first step. □

Step 2 of the Proof of Theorem 3.5: Using the learning error bound in (15), we obtain an upper bound on the probability of learning error exceeding a fixed $\delta > 0$ as shown below

$$\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) \leq \mathbb{T}_1 + \mathbb{T}_2 \left(\frac{\delta}{2 b_z} \right) + \mathbb{T}_5 \left(\frac{\delta}{2} \right), \tag{16}$$

where $\mathbb{T}_5(\delta) := (1 - q_w(\delta))^T$.

First, using the the fact that $\{w_t\}_{t=0}^{T-1}$ are iid, we show that

$$\begin{aligned}
\mathbb{P}\left(w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta \right) &= \mathbb{P}(w_{\max} - \delta > \|w_t\|_\infty, \ \forall\, 0 \leq t \leq T-1), \\
&= \prod_{t=0}^{T-1} \mathbb{P}(w_{\max} - \delta > \|w_t\|_\infty), \\
&\leq \prod_{t=0}^{T-1} \mathbb{P}(w_{\max} - \delta > w_t^1), \\
&\leq (1 - q_w(\delta))^T,
\end{aligned}$$

where the first inequality is due to $w_t^1 \leq \|w_t\|_\infty$ and the second inequality is from Assumption 2.6. Finally, we obtain the desired convergence rate using the error bound in (15) as follows

$$\mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta) \leq \mathbb{P}\left(b_z \text{diam}(\Theta_T) + w_{\max} - \max_{0 \leq t \leq T-1} \|w_t\|_\infty > \delta \right)$$

$$
\begin{aligned}
&\leq \mathbb{P}\left(b_z \mathrm{diam}(\Theta_T) > \delta/2 \text{ or } w_{\max} - \max_{0\leq t\leq T-1} \|w_t\|_\infty > \delta/2\right) \\
&\leq \mathbb{P}\left(\mathrm{diam}(\Theta_T) > \frac{\delta}{2b_z}\right) + \mathbb{P}\left(w_{\max} - \max_{0\leq t\leq T-1} \|w_t\|_\infty > \delta/2\right) \\
&\leq \mathbb{T}_1 + \mathbb{T}_2\left(\frac{\delta}{2b_z}\right) + \mathbb{T}_5(\delta/2).
\end{aligned}
$$

where the last inequality is by Theorem 3.1.

This completes the second and the last step of the proof. □

H.2. Proof of Corollary 3.6

First, by the proof of Corollary 3.2 in Appendix E, we have that $\mathbb{T}_1 = \frac{T}{m}\tilde{O}(n_z^{2.5})a_2^{n_z}\exp(-a_3 m) \leq \epsilon$ whenever $m \geq O(n_z + \log T + \log\frac{1}{\epsilon})$.

Next, we show $\mathbb{T}_5(\delta_T/2) \leq \mathbb{T}_2(\frac{\delta_T}{2b_z})$. Since $b_z \geq \sigma_z$ by the definition of BMSB, we have $\frac{a_1\delta_T}{8\sqrt{n_x}b_z} \leq \frac{\delta_T}{2}$. Since $q_w(\cdot)$ is a non-decreasing function, we have $1 - q_w(\frac{a_1\delta_T}{8\sqrt{n_x}b_z}) \geq 1 - q_w(\frac{\delta_T}{2})$. Notice that $m \geq 1$, and the constant factors in front of the $(1-q_w(\cdot))^{\lceil (T/m)\rceil}$ in $\mathbb{T}_2$ is also larger than 1. Consequently, $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \geq \mathbb{T}_5(\delta/2)$. Therefore, the choice of δ_T for the second term $\mathbb{T}_2$ also guarantees $\mathbb{T}_5(\delta_T/2) \leq \epsilon$.

Therefore, it suffices to ensure $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$. Notice that, when $\frac{\delta_T}{2b_z} = 2w_{\max}$, then $\mathbb{T}_2(\frac{\delta_T}{2b_z}) = 0 \leq \epsilon$, so there exists δ_T such that $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$.

Next, we will show that there exists such δ_T that diminishes to zero as T goes to infinity. Notice that we need

$$
1 - q_w\left(\frac{a_1\delta_T}{8b_z\sqrt{n_x}}\right) \leq \left(\frac{\epsilon}{\tilde{O}((n_x n_z)^{2.5} a_4^{n_x n_z})}\right)^{1/\lceil T/m\rceil},
$$

so that

$$
q_w\left(\frac{a_1\delta_T}{8b_z\sqrt{n_x}}\right) \geq 1 - \left(\frac{\epsilon}{\tilde{O}((n_x n_z)^{2.5} a_4^{n_x n_z})}\right)^{1/\lceil T/m\rceil},
$$

where the right hand side converges to zero as $T \to \infty$.

Now, consider $\delta(k) = 1/k$. Since $q_w\left(\frac{a_1\delta(k)}{8b_z\sqrt{n_x}}\right) > 0$, there exists a large enough T_k for any $k > 0$ such that for any $T \geq T_k$, we have that

$$
q_w\left(\frac{a_1\delta(k)}{8b_z\sqrt{n_x}}\right) \geq 1 - \left(\frac{\epsilon}{\tilde{O}((n_x n_z)^{2.5} a_4^{n_x n_z})}\right)^{1/\lceil T_k/m\rceil}.
$$

Furthermore, for any $T > 0$, we can define

$$
\delta_T = \begin{cases} \delta(k), & \text{if } T_k \leq T < T_{k+1}, \text{ for } k > 0, \\ 2w_{\max}, & \text{if } T < T_1. \end{cases}
$$

In this way, δ_T satisfies $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$ and $\delta_T \to 0$ as $T \to +\infty$.

Finally, using the proof of Corollary 3.2, we can show that there exists $\frac{\delta_T}{2b_z} = \tilde{O}(n_x^{1.5}n_z^2/T)$ such that $\mathbb{T}_2(\frac{\delta_T}{2b_z}) \leq \epsilon$ whenever $q_w(\delta) = O(\delta)$. This implies $\delta_T = \tilde{O}(n_x^{1.5}n_z^2/T)$ and completes the proof. □

H.3. Proof of Theorem 3.7

We first show that the unknown θ^* is a member of USC-SME uncertainty set $\hat{\Theta}_T^{\text{ucb}}$ with high probability. By Theorem 3.5, Corollary 3.6, and the definition in (5), we have

$$
\mathbb{P}(w_{\max} > \hat{w}_{\max}^{(T)}) = \mathbb{P}(w_{\max} - \bar{w}_{\max}^{(T)} > \delta_T) \leq 3\epsilon,
$$

which implies $1 - 3\epsilon \leq \mathbb{P}(w_{\max} \leq \hat{w}_{\max}^{(T)}) \leq \mathbb{P}(\theta^* \in \hat{\Theta}_T^{\text{ucb}})$.

Next, we show that the diameter of the UCB-SME uncertainty set is controlled by δ_T with high probability. Notice that $\hat{\Theta}_T^{\text{ucb}} \subseteq \Theta_T(w_{\max} + \delta_T)$ because $\bar{w}_{\max}^{(T)} \leq w_{\max}$. Therefore, by Theorem 3.4, the following holds for any constant $r > 0$:

$$\begin{aligned}\mathbb{P}(\text{diam}(\hat{\Theta}_T^{\text{ucb}}) > r + a_5\sqrt{n_x}\delta_T) &\leq \mathbb{P}(\text{diam}(\Theta_T(w_{\max} + \delta_T)) > r + a_5\sqrt{n_x}\delta_T),\\ &\leq \mathbb{T}_1 + \mathbb{T}_2(r).\end{aligned}$$

Let $r = \delta_T$, then, using the inequality $\mathbb{T}_2(\delta_T) \leq \mathbb{T}_2(\delta_T/2b_z)$, we have that

$$\begin{aligned}\mathbb{P}(\text{diam}(\hat{\Theta}_T^{\text{ucb}}) > \delta_T + a_5\sqrt{n_x}\delta_T) &\leq \mathbb{P}(\text{diam}(\Theta_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T),\\ &\leq 2\epsilon.\end{aligned}$$

Therefore, with probability $1 - 2\epsilon$, the diameter of $\hat{\Theta}_T^{\text{ucb}}$ is bounded above by

$$\text{diam}(\hat{\Theta}_T^{\text{ucb}}) \leq \delta_T + a_5\sqrt{n_x}\delta_T = O(\sqrt{n_x}\delta_T).$$

Finally, we can verify that the event $\{\text{diam}(\hat{\Theta}_T^{\text{ucb}}) \leq \delta_T + a_5\sqrt{n_x}\delta_T = O(\sqrt{n_x}\delta_T)\}$ and the event $\{\theta^* \in \hat{\Theta}_T^{\text{ucb}}\}$ simultaneously happen with probability at least $1 - 3\epsilon$ as follows:

$$\begin{aligned}&\mathbb{P}\left(\theta^* \notin \hat{\Theta}_T(\hat{w}_{\max}^{(T)}), \;\text{ or }\; \text{diam}(\hat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T\right)\\ &\leq \mathbb{P}\left(w_{\max} - \max_{0\leq t\leq T-1}\|w_t\|_\infty \geq \delta_T/2, \;\text{ or }\; \text{diam}(\Theta_T) > \delta_T/2b_z, \;\text{ or }\; \text{diam}(\hat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T\right)\\ &\leq \mathbb{P}\left(w_{\max} - \max_{0\leq t\leq T-1}\|w_t\|_\infty \geq \delta_T/2\right) + \mathbb{P}\left(\text{diam}(\Theta_T) > \delta_T/2b_z, \;\text{ or }\; \text{diam}(\hat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T\right)\\ &\leq \epsilon + \mathbb{P}(\mathcal{E}_2) + \sum_{i=1}^{v_\gamma}\mathbb{P}\left(\bigcap_k G_{i,k}(\min(\delta_T/2b_z, \delta_T))\right)\\ &\leq 3\epsilon.\end{aligned}$$

The third inequality follows from

- the proof of Theorem 3.5 in Appendix H.1,
- Theorem 3.4,
- the fact that the probabilties $\mathbb{P}(\text{diam}(\Theta_T) > \delta_T/2b_z)$ and $\mathbb{P}(\text{diam}(\hat{\Theta}_T(w_{\max} + \delta_T)) > \delta_T + a_5\sqrt{n_x}\delta_T)$ are bounded by the same events $\mathcal{E}_2$,
- and $G_{i,k}(\delta_T), G_{i,k}(\delta_T/2b_z) \subseteq G_{i,k}(\min(\delta_T/2b_z, \delta_T))$, where $G_{i,k}(\delta)$ is defined in Lemma 4.4 as a function of δ.

This completes the proof. □

I. Simulation details and additional experiments

This section provides the details on the simulation experiments, along with some additional results. The code for replicating the presented results can be found in the github repository: `https://github.com/jy-cds/non-asymptotic-set-membership`.

I.1. Baseline: LSE's confidence regions

In all our experiments, we use the 90% confidence region of the LSE as the baseline uncertainty set. The diameters of LSE's confidence regions are computed by taking minimum of the formulas provided in the following two papers: Lemma E.3 in (Simchowitz & Foster, 2020) and Theorem 1 in (Abbasi-Yadkori & Szepesvári, 2011). To apply Theorem 1 in (Abbasi-Yadkori & Szepesvári, 2011), we used regularization parameter $\lambda = 0.1$, $\delta = 0.1$ for 90% confidence, $S = \sqrt{\text{tr}\,(\theta^{*,\top}\theta^*)}$, variance proxy $L = 1$ for truncated Gaussian distribution and $L = 4/3$ for uniform distribution.

To determine the parameters in Lemma E.3 of (Simchowitz & Foster, 2020), we approximately optimize the projection matrix P in Lemma E.3 as follows. First, we consider an orthogonal transformation of the empirical covariance matrix $\Lambda = \sum_{t=1}^{T} z_t z_t^\top$ with $\Lambda = GMG^\top$ where G is unitary. This transforms the event $\mathcal{E}$ in Lemma E.3 to $M \geq \lambda_1 P_0 + \lambda_2(I - P_0)$, where $GP_0G^\top = P$. We select P_0 as a block matrix $[[I_p,\, 0], [0,\, 0]]$, then optimize over the block size p in search of the tightest LSE confidence bound.

I.2. Figure 1: SME and LSE uncertainty set visualization

In this experiment, we consider $x_{t+1} = A^* x_t + B^* u_t + w_t$, where $A^* = 0.8$ and $B^* = 1$ are unknown. $w_t \sim \text{TrunGauss}(0, \sigma_w, [-w_{\max}, w_{\max}])$ is i.i.d. and $u_t \sim \text{TrunGauss}(0, \sigma_u, [-u_{\max}, u_{\max}])$ are also i.i.d generated, where $\sigma_w = \sigma_u = 0.5$, and $w_{\max} = u_{\max} = 1$. We compare SME that knows $w_{\max} = 1$ and LSE's 90% confidence region computed based on Appendix I.1.

I.3. Figure 2

In this experiment, we consider the the linearized longitudinal flight control dynamics of Boeing 747 (Lale et al., 2022; Mete et al., 2022) with i.i.d. bounded inputs and disturbances sampled from truncated Gaussian and uniform distribution. The dynamics is $x_{t+1} = A^* x_t + B^* u_t + w_t$ with

$$A = \begin{bmatrix} 0.99 & 0.03 & -0.02 & -0.32 \\ 0.01 & 0.47 & 4.7 & 0 \\ 0.02 & -0.06 & 0.4 & 0 \\ 0.01 & -0.04 & 0.72 & 0.99 \end{bmatrix} \quad B = \begin{bmatrix} 0.01 & 0.99 \\ -3.44 & 1.66 \\ -0.83 & 0.44 \\ -0.47 & 0.25 \end{bmatrix}.$$

Disturbances are sampled from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^4)$ as well as $\text{Unif}([-w_{\max}, w_{\max}]^4)$, while control inputs are samples from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^2)$ in both disturbance settings, with $w_{\max} = 2$. To compute the UCB for SME using (5), we heuristically define $\delta_T = \beta \, \frac{n_x^{1.5} \cdot n_z^2 \cdot (\max_t \|x_t\|)}{T}$, where $n_x = 4$ and $n_z = 6$ are the system dimension, while β is a tunable parameter. This definition matches the dimension and time order of the theoretical analysis in Corollary 3.6. In both experiments of Figure 2, we fix $\beta = 0.01$.

In Figure 2(a)-(b), we plot SME with accurate and conservative bounds of $w_{\max}$, UCB-SME, and LSE's 90% confidence regions computed by Appendix I.1. We use 10 different seeds to generate the disturbance sequences for each plot, and use the shaded region to denote 1 standard deviation from the mean (colored lines).

I.4. Figure 3

In this experiment, we consider autonomous systems of the form $x_{t+1} = A^* x_t + w_t$, where $A^* \in \mathbb{R}^{n_x}$ is randomly sampled and its spectral radius is normalized to be 0.9. We simulate SME and LSE for $n_x = 5, 10, 15, 20, 25$. The disturbances are sampled from $\text{TrunGauss}(0, I, [-w_{\max}, w_{\max}]^{n_x})$ as well as $\text{Unif}([-w_{\max}, w_{\max}]^{n_x})$ with $w_{\max} = 2$. This simula!tion is run on 10 random seeds and the total length of the simulation is set to be $T = 1000$ across all n_x experiments. The mean is plotted as solid lines and the shaded regions denote 1 standard deviation from the mean.

Though SME's theoretical bound with respect to the dimension is $\tilde{O}(n_x^{1.5} n_z^2)$ from Corollary 3.3, which is much worse than LSE's bound, it is not reflected in Figure 3. Therefore, it is promising that the dimension scaling in the analysis in Section 3 can be further tightened. We leave this for future work.

I.5. Figure 4

To illustrate the quantitative impact of using SME for adaptive tube-based robust MPC, we study tube-based robust MPC for a system $x_{t+1} = A^* x_t + B^* u_t + w_t$ with nominal system $A^* = 1.2$, $B^* = 0.9$ with an initial model uncertainty set $\Theta_0 := [1, 1.2] \times [0.9, 1.1]$. We use the basic tube-based robust MPC method (Rawlings et al., 2017; Mayne et al., 2005) and parameterize the control policy as $u_k = K x_k + v_k + \eta_k$, where $K = -1$, v_k is determined by the tube-based robust MPC algorithm, and η_k is a bounded exploration injection with $\eta_k \sim \text{Unif}([-0.01, 0.01])$. The disturbance w_k has a known bound of $w_{\max} = 0.1$ and is generated to be i.i.d. $\text{Unif}([-0.1, 0.1])$. The horizon of the tube-based robust MPC is set to be 5. The state and input constraints are such that $x_k \in [-10, 10]$ and $u_k \in [-10, 10]$ for all $k \geq 0$. We consider the task of constrained LQ tracking problem with a time-varying cost function $c_t := (x_t - g_t)^\top Q (x_t - g_t) + u_t^\top R u_t$ where the target trajectory is generated as $g_t = 8 \sin(t/20)$.

We compare the performance of an adaptive tube-based robust MPC controller that uses the SME for uncertainty set estimation against one that uses the LSE 90% confidence region (LSE). For better visualization of the trajectory difference as a result of different estimation methods, we used the minimum of the the dominant factors in Dean et al. (2018, equation C.12) and the LSE 90% confidence region for the LSE uncertainty set. We also plot the offline optimal RMPC controller, i.e., the controller that has knowledge of the true underlying system parameters (OPT).

Since the controller has to robustly satisfy constraints against the worst-case model in the uncertainty set, smaller uncertainty set for the tube-based robust MPC means more optimal trajectories can be computed. This observation is consistent with the extensive empirical results in the control literature (Lorenzen et al., 2019; Lu et al., 2019; Köhler et al., 2019).