

# Control of Discrete-Time LTI Systems Using Stochastic Ensemble Systems

Nirabhra Mandal

Mohammad Khajenejad

Sonia Martínez

## Abstract

In this paper, we study the control properties of a new class of stochastic ensemble systems that consists of families of random variables. These random variables provide an increasingly good approximation of an unknown discrete, linear-time invariant (DLTI) system, and can be obtained by a standard, data-driven procedure. Our first result relates the reachability properties of the stochastic ensemble system with that of the limiting DLTI system. We then provide a method to combine the control inputs obtained from the stochastic ensemble systems to compute a control input for the DLTI system. Later, we deal with a particular kind of stochastic ensemble systems generated from realizing Bernoulli random variables. For this, we characterize the variance of the computed state and control. We also do the same for a situation where the data is updated sequentially in a streaming fashion. We illustrate the results numerically in various simulation examples.

**Keywords** — Approximate reachability, sample reachability, stochastic ensemble system.

## 1 Introduction

Ensemble control, *i.e.*, investigating the ability of steering and manipulating an entire ensemble of (partially) unknown systems in a desired and optimal manner, has emerged from several science and engineering applications in recent years, *e.g.* coordination of the movement of flocks in biology [1], manipulation of spin ensembles in nuclear magnetic resonance [2, 3], or desynchronization of pathological neurons in the brain in neuroscience [4].

**Literature Review** Motivated by this, a huge body of seminal work has been done on the analysis of (deterministic) ensemble controllability [5, 6, 7] and synthesis of optimal ensemble controls [8, 9, 10] by developing new analytical and numerical methods, which have received increasing attention due to their application to robotics [7], energy systems [11, 12] and quantum control [13, 14]. However, the traditional ensemble control problem consists of driving a collection of initial states of a continuum of systems to a set of final states with the same control input [15], which could be potentially conservative. The problem of reachability for finite-dimensional linear time-varying systems has been studied in [16, 17], but here the system varies in

a continuum and not in a countable set. On the other hand, the reachability properties of bilinear ensemble systems are studied in [18]. In [8], the authors explore the problem of optimal control for stochastic linear systems. The stochastic nature comes from additive noise to the dynamics. In [19, 20] the authors study robustness of matrix properties to unknown system parameters. Reference [11] solves an ensemble control problem for devices with cyclic energy consumption patterns by utilizing techniques from the Markov Decision Process framework. Finally, the work [12] looks into a similar problem but with uncertain dynamics and uses techniques from stochastic and distributionally robust optimization in the process. However, in all the aforementioned work, either the control signal needs to be unique, or a continuum of ensemble (and partially unknown) systems is required to exist. This work aims to bridge this gap.

**Contributions** Unlike ensemble control, we let the system parameters take values in a countable set. Moreover, we do not restrict the control to be unique but allow every sample run of the system to employ a different control function. This being said, we consider a so-called class of stochastic ensemble systems, which arise from the approximation of systems whose exact parameters are unknown. A known approach to obtain good parametric models consists of learning the distribution of for the model whose realizations correspond to a system approximation. As more data becomes available, the distributions become more accurate, resulting in a stochastic ensemble approximation. This enables us to i) study the reachability properties of a new class of stochastic ensemble systems, ii) provide a procedure for combining controls from such systems, iii) compute a control for the limiting DLTI system, iv) characterize the variance of the state and control of a stochastic ensemble system produced by means of a Bernoulli distribution, and v) derive an improved result by means of least squares error minimization. As conclusions of these main contributions, we also can vi) compute the variance of the state and control of a stochastic ensemble system that is obtained in a streaming fashion. We illustrate our results in numerical examples.

**Notations** We denote the set of real numbers, the set of non-negative real numbers, the set of integers,

\*Nirabhra Mandal, Mohammad Khajenejad and Sonia Martínez are with the Mechanical and Aerospace Engineering Department, University of California San Diego. {nmandal,mkhajenejad,soniamd}@ucsd.edu. This work was partially supported by AI2C ARL, Grant No: W911NF-23-2-0009.

and the set of non-negative integers using  $\mathbb{R}$ ,  $\mathbb{R}_{\geq 0}$ ,  $\mathbb{Z}$ , and  $\mathbb{Z}_{\geq 0}$ , respectively. We let  $\mathbb{R}^n$  (similarly  $\mathbb{Z}_{\geq 0}^n$ ) be the Cartesian product of  $\mathbb{R}$  (similarly  $\mathbb{Z}_{\geq 0}$ ) with itself.  $\mathbb{R}^{n \times m}$  denotes the set of real matrices of order  $n \times m$ . The  $i^{\text{th}}$  component of a vector  $\mathbf{v} \in \mathbb{R}^n$  is denoted by  $\mathbf{v}_i$  and the  $ij^{\text{th}}$  entry of a matrix  $\mathbf{M} \in \mathbb{R}^{n \times m}$  is denoted by  $\mathbf{M}_{ij}$ . For a set of square matrices  $\{\mathbf{M}^{(s)}\}_{s \in \mathbb{Z}}$ , the product operator is used to denote multiplication from right to left, i.e.  $\prod_{i=1}^n \mathbf{M}^{(s)} := \mathbf{M}^{(n)} \cdots \mathbf{M}^{(1)}$ . For a vector  $\mathbf{v} \in \mathbb{R}^n$ ,  $\mathbf{v} \geq 0$  denotes a component-wise inequality. We let  $\text{abs} : \mathbb{R}^n \rightarrow \mathbb{R}^n$  be a function that produces a vector with component-wise absolute values from the input vector, i.e.  $\text{abs}(\mathbf{v})_i = |\mathbf{v}_i| \in \mathbb{R}$ ,  $\forall i \in \{1, \dots, n\}$ . We denote the empty set using  $\emptyset$ .

## 2 Problem Formulation

Consider a discrete-time, unknown, linear time-invariant (LTI) system of the form

$$(2.1) \quad \mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k),$$

where  $\mathbf{x} \in \mathbb{R}^n$  is the state,  $\mathbf{x}(0) = \mathbf{x}_0$  and  $\mathbf{x}_f$  are the initial and desired final states, respectively,  $\mathbf{u} \in \mathbb{R}^m$  is the control input,  $\mathbf{A} \in \mathbb{R}^{n \times n}$  and  $\mathbf{B} \in \mathbb{R}^{n \times m}$ . The system matrices  $\mathbf{A}$  and  $\mathbf{B}$ , as well as the initial state  $\mathbf{x}_0$  are unknown, but it is possible to construct increasingly good approximations of these system parameters using some known approach, e.g., any data-driven or Machine Learning-based method. This approximation is assumed to be done through the following newly defined class of *stochastic ensemble systems*, for which we aim to characterize reachability properties.

**DEFINITION 2.1.** (*Stochastic ensemble system*). Consider sequences of independent random variables  $\mathcal{X}_0 := \{\mathbf{x}_0^{(s)} \in \mathbb{R}^n\}_{s \in \mathbb{Z}_{\geq 0}}$ ,  $\mathcal{A} := \{\mathbf{A}^{(s)} \in \mathbb{R}^{n \times n}\}_{s \in \mathbb{Z}_{\geq 0}}$  and  $\mathcal{B} := \{\mathbf{B}^{(s)} \in \mathbb{R}^{n \times m}\}_{s \in \mathbb{Z}_{\geq 0}}$  such that  $\mathcal{A}$ ,  $\mathcal{B}$ , and  $\mathcal{X}_0$  are all independent of each other. For every  $\sigma \in \mathbb{Z}_{\geq 0}$ , a stochastic ensemble system is a linear, time varying (LTV) system of the form:

$$(2.2a) \quad \widehat{\mathbf{x}}(\sigma, k+1) = \widehat{\mathbf{A}}(s_A(\sigma, k)) \widehat{\mathbf{x}}(\sigma, k) + \widehat{\mathbf{B}}(s_B(\sigma, k)) \widehat{\mathbf{u}}(\sigma, k)$$

$$(2.2b) \quad \widehat{\mathbf{x}}(\sigma, 0) = \mathbf{x}_0^{(s_0(\sigma))} \in \mathcal{X}_0,$$

where, the indicator functions  $s_A : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$ ,  $s_B : \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  and  $s_0 : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  are used to index different realizations of the random variables in  $\mathcal{X}_0$ ,  $\mathcal{A}$ , and  $\mathcal{B}$ , respectively. Thus,  $\widehat{\mathbf{A}}(s_A(\sigma, k))$  (similarly  $\widehat{\mathbf{B}}(s_B(\sigma, k))$ ) represents a realization of  $\mathbf{A}^{(s_A(\sigma, k))} \in \mathcal{A}$  (respectively  $\mathbf{B}^{(s_B(\sigma, k))} \in \mathcal{B}$ ) for each  $\sigma$  and  $k$ . •

We use the terms random variable and realization interchangeably. The  $\sigma$  parameter is used to distinguish

between different runs of the system. Also note that since the initial state and the system matrices in (2.2) are realizations of different random variables, each state  $\widehat{\mathbf{x}}(\sigma, k)$  is also a realization of some random variable. Moreover, we need to formally define the notion of *stochastic ensemble approximation*.

**DEFINITION 2.2.** (*Stochastic ensemble approximation*). Let  $\mathbf{M} \in \mathbb{R}^{n \times m}$ . Consider a sequence of independent random variables  $\mathcal{M} := \{\mathbf{M}^{(s)}\}_{s \in \mathbb{Z}_{\geq 0}}$  such that  $\exists B \in \mathbb{R}$  such that  $\|\mathbf{M}^{(s)}\| \leq B$ ,  $\forall s \in \mathbb{Z}_{\geq 0}$ . Then,  $\mathcal{M}$  is a stochastic ensemble approximation of  $\mathbf{M}$  if and only if

$$\forall \varepsilon > 0, \mathbb{P}(\|\mathbf{M} - \sum_{s=1}^d w_s \mathbf{M}^{(s)}\|_p > \varepsilon) \rightarrow 0 \text{ as } d \rightarrow \infty,$$

with respect to some  $p$ -norm  $\|\cdot\|_p$  and for some  $w_s$ 's of the form  $[w_1, \dots, w_d]^T \in \mathcal{S}^d := \{\mathbf{y} \in \mathbb{R}^d \mid \mathbf{1}^T \mathbf{y} = 1, \mathbf{y} \geq 0\}$  for each  $d \in \mathbb{Z}_{\geq 0}$ . •

Furthermore, we assume the following.

**ASSUMPTION 2.1.** There exists a stochastic ensemble system of the form (2.2) that consists of stochastic ensemble approximations of the original system given in (2.1). •

Being concerned with meeting certain reachability requirements for (2.2), next we provide a formal definition for the notion of approximate reachability.

**DEFINITION 2.3.** (*Approximate reachability of final state from an initial state*). Let  $\mathcal{X}_0$ ,  $\mathcal{A}$ , and  $\mathcal{B}$  be stochastic ensemble approximations of  $\mathbf{x}_0$ ,  $\mathbf{A}$ , and  $\mathbf{B}$  respectively. Consider the stochastic ensemble system in (2.2) and let  $\mathbf{x}_f \in \mathbb{R}^n$ . Suppose there exists a fixed time  $K \in \mathbb{Z}_{\geq 0}$  and for any realization of the random variables  $\widehat{\mathbf{x}}(\sigma, 0)$ ,  $\widehat{\mathbf{A}}(s_A(\sigma, k))$ , and  $\widehat{\mathbf{B}}(s_B(\sigma, k))$  there exist controls  $\widehat{\mathbf{u}}(\sigma, 0), \dots, \widehat{\mathbf{u}}(\sigma, K-1)$  that drive the realization  $\widehat{\mathbf{x}}(\sigma, 0)$  through the trajectory  $\widehat{\mathbf{x}}(\sigma, k)$ ,  $\forall k \in \{1, \dots, K\}$  under the dynamics (2.2), i.e.

$$(2.3) \quad \begin{aligned} \widehat{\mathbf{x}}(\sigma, K) &= \prod_{k=0}^{K-1} \widehat{\mathbf{A}}(s_A(\sigma, k)) \widehat{\mathbf{x}}(\sigma, 0) \\ &+ \sum_{k=0}^{K-1} [\prod_{i=k+1}^{K-1} \widehat{\mathbf{A}}(s_A(\sigma, i))] \widehat{\mathbf{B}}(s_B(\sigma, k)) \widehat{\mathbf{u}}(\sigma, k). \end{aligned}$$

Then  $\mathbf{x}_f$  is said to be approximately reachable from  $\mathbf{x}_0$  if and only if  $\{\widehat{\mathbf{x}}(\sigma, K)\}_{\sigma \in \mathbb{Z}_{\geq 0}}$  is a stochastic ensemble approximation of  $\mathbf{x}_f$ . •

Further, we need to formally introduce the notion of sample reachability as follows.

**DEFINITION 2.4.** (*Sample reachability of final state from an initial state in unit time*). Let  $\mathcal{X}_0$ ,  $\mathcal{X}_f$ , and  $\mathcal{A}$  be stochastic ensemble approximations of  $\mathbf{x}_0$ ,  $\mathbf{x}_f$  and  $\mathbf{A}$  respectively, and  $s_f : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{Z}_{\geq 0}$  be an indicator function that denotes the realizations of the random variables in

the set  $\mathcal{X}_f := \{\mathbf{x}_f^{(s)} \in \mathbb{R}^n\}_{s \in \mathbb{Z}_{\geq 0}}$ . Consider the stochastic ensemble system in (2.2) with  $\widehat{\mathbf{B}}(k) = \mathbf{B}$ . Suppose there exists a control  $\{\widehat{\mathbf{u}}(\sigma, 0)\}_{\sigma \in \mathbb{Z}_{\geq 0}}$  for each  $\sigma \in \mathbb{Z}_{\geq 0}$  that drives  $\widehat{\mathbf{x}}(\sigma, 0) = \mathbf{x}_0^{(s_0(\sigma))}$  to  $\widehat{\mathbf{x}}(\sigma, K) = \mathbf{x}_f^{(s_f(\sigma))}$  under the dynamics (2.2), in one ( $K = 1$ ) time step, i.e.

$$(2.4) \quad \mathbf{x}_f^{(s_f(\sigma))} = \widehat{\mathbf{A}}(s_A(\sigma, 1)) \mathbf{x}_0^{(s_0(\sigma))} + \mathbf{B} \widehat{\mathbf{u}}(\sigma, 0).$$

Then  $\mathbf{x}_f$  is sample reachable from  $\mathbf{x}_0$  in unit time. •

Note that the control  $\widehat{\mathbf{u}}(\sigma, 0)$  is chosen in order to drive  $\mathbf{x}_0^{(s_0(\sigma))}$  to  $\mathbf{x}_f^{(s_f(\sigma))}$  in the previous definition.

Our problem of interest can be cast as follows.

**PROBLEM 2.1.** *Given the setup and Assumption 2.1, characterize the relation between approximate reachability under the stochastic ensemble system (2.2) and the standard reachability under the DLTI system (2.1). Moreover, find controls that drive  $\mathbf{x}_0$  to  $\mathbf{x}_f$  under (2.1) using the stochastic ensemble system (2.2).* •

We conclude this section by highlighting the difference between our notion of stochastic ensemble system with that of (deterministic) ensemble control.

**REMARK 2.1.** (*Comparison with ensemble control*). The stochastic ensemble system in (2.2) has a similar structure to ensemble systems [5]. However, the  $\sigma$  that parameterizes the system takes discrete values. Moreover,  $\widehat{\mathbf{A}}$  and  $\widehat{\mathbf{B}}$  do not vary continuously with respect to  $\sigma$ . We also allow the controls for each  $\sigma$  to be different, i.e., we do not seek one single control that steers the ensemble system between points of interest. Finally, if a system is ensemble controllable, then similar results to the ones in this paper can be given, generalizing ensemble control systems and their objective. •

### 3 Control using Stochastic Ensemble Systems

In this section, we first show that reachability under the DLTI system (2.1) is sufficient for approximate reachability<sup>1</sup> under the stochastic ensemble systems.

Recall that we use  $\sigma \in \mathbb{Z}_{\geq 0}$  to index the runs of the system. Moreover,  $s_0(\sigma)$  (respectively  $s_A(\sigma, k)$ ,  $s_B(\sigma, k)$ ) indexes the realization of the initial state (respectively of matrices  $\mathbf{A}$  and  $\mathbf{B}$  at the  $k^{\text{th}}$  time step) used in (2.2). In the sequel, we omit the  $\sigma$  for the sake of brevity wherever there is no confusion. Now we are ready to state the result.

**LEMMA 3.1.** (*Sufficient condition for approximate reachability using state reachability*). Let  $\mathcal{X}_0$ ,  $\mathcal{A}$ , and

<sup>1</sup>Not to be misinterpreted as the reachability of autonomous systems, studied e.g., in the authors' previous works [21, 22].

$\mathcal{B}$  be stochastic ensemble approximations of  $\mathbf{x}_0$ ,  $\mathbf{A}$ , and  $\mathbf{B}$  respectively. Suppose that  $\mathbf{x}_f$  is reachable from  $\mathbf{x}_0$  in  $K \in \mathbb{Z}_{\geq 0}$  time steps under the dynamics (2.1). Then,  $\mathbf{x}_f$  is approximately reachable from  $\mathbf{x}_0$  under the stochastic ensemble system in (2.2).

*Proof.* Since  $\mathcal{X}_0$ ,  $\mathcal{A}$ , and  $\mathcal{B}$  are stochastic ensemble approximations of  $\mathbf{x}_0$ ,  $\mathbf{A}$ , and  $\mathbf{B}$  respectively  $\exists B_{\mathcal{X}_0}, B_{\mathcal{A}}$ , and  $B_{\mathcal{B}}$  such that  $\|\mathbf{x}_0^{(s)}\| \leq B_{\mathcal{X}_0}$ ,  $\|\mathbf{A}^{(s)}\| \leq B_{\mathcal{A}}$ , and  $\|\mathbf{B}^{(s)}\| \leq B_{\mathcal{B}}$ ,  $\forall s \in \mathbb{Z}_{\geq 0}$ . Now, by the hypothesis, there exists controls  $\mathbf{u}(0), \dots, \mathbf{u}(K-1)$  that drive  $\mathbf{x}_0$  to  $\mathbf{x}_f$  under the dynamics (2.1). Moreover since  $K$  is finite, all the controls can be bounded by a constant  $B_{\mathbf{u}}$ . We show that  $\widehat{\mathbf{u}}(\sigma, k) = \mathbf{u}(k)$ ,  $\forall k \in \{1, \dots, K-1\}$  proves the claim. Let  $\varepsilon > 0$ . If we use the controls  $\mathbf{u}(0), \dots, \mathbf{u}(K-1)$  in (2.3), we obtain

$$\mathbf{y}(1) - \sum_{s_0=1}^N w_{s_0,1} \widehat{\mathbf{x}}(s(1)) = \widehat{\mathbf{A}}(s_A(1)) (\mathbf{x}_0 - \sum_{s_0=1}^N w_{s_0,1} \mathbf{x}_0^{(s_0)}) + \widehat{\mathbf{B}}(s_B(1)) (\mathbf{u}(0) - \sum_{s_0=1}^N w_{s_0,1} \mathbf{u}(0)),$$

at  $k = 1$  starting at  $\mathbf{x}_0^{(s_0)}$  and for  $[w_{s_0,1}] = [w_{1,1}, \dots, w_{N,1}] \in \mathcal{S}^N$  which makes  $\mathcal{X}_0$  a stochastic ensemble approximation of  $\mathbf{x}_0$  and  $\mathbf{y}(1) = \widehat{\mathbf{A}}(s_A(\sigma, 1)) \mathbf{x}_0 + \widehat{\mathbf{B}}(s_B(\sigma)) \mathbf{u}(0)$ . Note that here we are using the weights that make  $\mathcal{X}_0$  a stochastic ensemble approximation throughout and we use the independence properties of the random variables to separate the product and the sum. Now, since  $\sum_{s_0=1}^N w_{s_0,1} \mathbf{u}(0) = \mathbf{u}(0)$ , then

$$(3.5) \quad \|\mathbf{y}(1) - \sum_{s_0=1}^N w_{s_0,1} \widehat{\mathbf{x}}(s(1))\| \leq B_{\mathcal{A}} \|\mathbf{x}_0 - \sum_{s_0=1}^N w_{s_0,1} \mathbf{x}_0^{(s_0)}\|.$$

Next, by defining  $\mathbf{z}(1) = \widehat{\mathbf{A}}(s_A(1)) \mathbf{x}_0 + \mathbf{B} \mathbf{u}(0)$  and considering a  $[w_{s_B(1),2}] = [w_{1,2}, \dots, w_{N,2}] \in \mathcal{S}^N$ ,  $\mathcal{B}$  becomes a stochastic ensemble approximation of  $\mathbf{B}$ . Then, by similar arguments as before

$$(3.6) \quad \|\mathbf{z}(1) - \sum_{s_B(1)=1}^N w_{s_B(1),2} \mathbf{y}(1)\| \leq B_{\mathbf{u}} \|\mathbf{B} - \sum_{s_B(1)=1}^N w_{s_B(1),2} \mathbf{B}(s_B(1))\|.$$

Finally, for  $\mathbf{x}(1) = \mathbf{A} \mathbf{x}_0 + \mathbf{B} \mathbf{u}(0)$ , we have

$$(3.7) \quad \|\mathbf{x}(1) - \sum_{s_A(1)=1}^N w_{s_A(1),3} \mathbf{z}(1)\| \leq \|\mathbf{A} - \sum_{s_A(1)=1}^N w_{s_A(1),3} \widehat{\mathbf{A}}(s_A(1))\| \|\mathbf{x}(0)\|,$$

using  $[w_{s_A(1),2}] = [w_{1,2}, \dots, w_{N,2}] \in \mathcal{S}^N$  which makes  $\mathcal{A}$  a stochastic ensemble approximation of  $\mathbf{A}$ . Then, using the triangle inequality of the norms and (3.5)–(3.7), we obtain

$$\|\mathbf{x}(1) - \sum_{s_0=1}^N w_{s_0,1} \widehat{\mathbf{x}}(\sigma, 1)\| \leq \alpha \|\mathbf{x}_0 - \sum_{s_0=1}^N w_{s_0,1} \mathbf{x}_0^{(s_0)}\| + \|\mathbf{B} - \sum_{s_B(1)=1}^N w_{s_B(1),2} \mathbf{B}(s_B(1))\| + \|\mathbf{A} - \sum_{s_A(1)=1}^N w_{s_A(1),3} \widehat{\mathbf{A}}(s_A(1))\|,$$

where  $\alpha = \max\{B_{\mathcal{A}}, B_{\mathbf{u}}, \|\mathbf{x}(0)\|\}$ . Hence,

$$\begin{aligned} & \mathbb{P} \left( \|\mathbf{x}(1) - \sum_{s_0=1}^N w_{s_0,1} \widehat{\mathbf{x}}(\sigma, 1)\| < \varepsilon \right) \geq \\ & \mathbb{P} \left( \|\mathbf{A} - \sum_{s_A(1)=1}^N w_{s_A(1),3} \widehat{\mathbf{A}}(s_A(1))\| < \frac{\varepsilon}{3\alpha} \right) + \\ & \mathbb{P} \left( \|\mathbf{B} - \sum_{s_B(1)=1}^N w_{s_B(1),2} \mathbf{B}(s_B(1))\| < \frac{\varepsilon}{3\alpha} \right) + \\ & \mathbb{P} \left( \|\mathbf{x}_0 - \sum_{s_0=1}^N w_{s_0,1} \mathbf{x}_0^{(s_0)}\| < \frac{\varepsilon}{3\alpha} \right) \rightarrow 1 \text{ as } N \rightarrow \infty. \end{aligned}$$

Thus,  $\{\widehat{\mathbf{x}}(s(1))\}_{s(1) \in \mathbb{Z}_{\geq 0}}$  forms a stochastic ensemble approximation of  $\mathbf{x}(1)$ . Thus, using the fact that  $\{\mathbf{x}_0^{(s_0)}\}_{s_0 \in \mathbb{Z}_{\geq 0}}$ ,  $\{\widehat{\mathbf{A}}(s_A(1))\}_{s_A(1) \in \mathbb{Z}_{\geq 0}}$ , and  $\{\widehat{\mathbf{B}}(s_B(1))\}_{s_B(1) \in \mathbb{Z}_{\geq 0}}$  are stochastic ensemble approximations of  $\mathbf{x}_0$ ,  $\mathbf{A}$ , and  $\mathbf{B}$  respectively, we have shown that  $\{\widehat{\mathbf{x}}(s(1))\}_{s(1) \in \mathbb{Z}_{\geq 0}}$  forms a stochastic ensemble approximation of  $\mathbf{x}(1)$ . Next using the fact that  $\{\widehat{\mathbf{x}}(s(1))\}_{s(1) \in \mathbb{Z}_{\geq 0}}$ ,  $\{\widehat{\mathbf{A}}(s_A(2))\}_{s_A(2) \in \mathbb{Z}_{\geq 0}}$ , and  $\{\widehat{\mathbf{B}}(s_B(2))\}_{s_B(2) \in \mathbb{Z}_{\geq 0}}$  are stochastic ensemble approximations of  $\mathbf{x}(1)$ ,  $\mathbf{A}$ , and  $\mathbf{B}$  respectively we can show that  $\{\widehat{\mathbf{x}}(s(2))\}_{s(2) \in \mathbb{Z}_{\geq 0}}$  forms a stochastic ensemble approximation of  $\mathbf{x}(2)$ . By induction on  $\mathbf{x}(1), \dots, \mathbf{x}(K)$ , the proof can be completed. ■

Note that the previous lemma states that the desired state  $\mathbf{x}_f$  is reachable from  $\mathbf{x}_0$  only if  $\mathbf{x}_f$  is approximately reachable from  $\mathbf{x}_0$ . Also, note that we deal with a much weaker version of reachability here. All we require is that  $\mathbf{x}_f$  be in the reachable subspace from  $\mathbf{x}_0$ . We do not require the whole  $\mathbb{R}^n$  to be reachable.

Next, by using the notion of sample reachability (Definition 2.4), we synthesize a control input sequence that drives the system in (2.1) from  $\mathbf{x}_0$  to  $\mathbf{x}_f$ .

**LEMMA 3.2.** (*Approximation of control of LTI systems using stochastic ensemble systems*). Consider the system in (2.1) with initial state  $\mathbf{x}_0$  and desired final state  $\mathbf{x}_f$ . Let  $\mathcal{X}_0, \mathcal{X}_f$  and  $\mathcal{A}$  be stochastic ensemble approximations of  $\mathbf{x}_0, \mathbf{x}_f$  and  $\mathbf{A}$  respectively. Next consider the stochastic ensemble system in (2.2) with  $\widehat{\mathbf{B}}(k) = \mathbf{B}$ ,  $\forall k \in \{1, \dots, K-1\}$ . Suppose that  $\mathbf{x}_f$  is sample reachable from  $\mathbf{x}_0$  in unit time. Let  $\{\widehat{\mathbf{u}}(\sigma, 0)\}_{\sigma \in \mathbb{Z}_{\geq 0}}$  be as in Definition 2.4. Then for each  $\varepsilon > 0$ ,

$$(3.8) \quad \lim_{N \rightarrow \infty} \mathbb{P}(\|\mathbf{B}\mathbf{v} - \mathbf{B} \sum_{\sigma=1}^N w_{\sigma} \widehat{\mathbf{u}}(\sigma, 0)\| > \varepsilon) \rightarrow 0$$

with  $w_{\sigma}$ 's of the form  $[w_1, \dots, w_N]^{\top} \in \mathcal{S}^N$  for each  $N \in \mathbb{Z}_{\geq 0}$  and with  $\mathbf{B}\mathbf{v}$  satisfying

$$(3.9) \quad \mathbf{x}_f = \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{v}.$$

*Proof.* We do this in a very similar way to the proof of Lemma 3.1. So we skip most of the details due to lack of space. Consider an  $\varepsilon > 0$ . Then, consider the state evolution equations

$$(3.10a) \quad \mathbf{x}_f^{(s_f)} = \widehat{\mathbf{A}}(s_A(1))\mathbf{x}_0 + \mathbf{B}\mathbf{u}'(0),$$

$$(3.10b) \quad \mathbf{x}_f = \widehat{\mathbf{A}}(s_A(1))\mathbf{x}_0 + \mathbf{B}\mathbf{u}''(0).$$

Note that (3.10a) has been formulated by choosing controls such that they drive the initial condition  $\mathbf{x}_0$  (and not the realizations) to the final state  $\mathbf{x}_f^{(s_f)}$ . Similarly (3.10b) has been formulated by choosing controls such that they drive the initial condition  $\mathbf{x}_0$  to the final state  $\mathbf{x}_f$  (and not the realizations). Then, using very similar arguments as in the proof of Lemma 3.1 it

is possible to show that the controls  $\mathbf{u}'$ , and  $\mathbf{u}''$  exist and can be attained by taking the weighted sum of the controls in (2.4).

Moreover, using the triangle inequality of the norms and probability theory, it can be shown that

$$\begin{aligned} \mathbb{P}(\|\mathbf{B}\mathbf{v} - \mathbf{B} \sum_{s_0=1}^N w_{s_0,1} \widehat{\mathbf{u}}(s_0)\| < \varepsilon) &\geq \\ \mathbb{P}(\|\mathbf{A} - \sum_{s_A(1)=1}^N w_{s_A(1),3} \widehat{\mathbf{A}}(s_A(1))\| < \frac{\varepsilon}{3\beta}) &+ \\ \mathbb{P}(\|\mathbf{x}_f - \sum_{s_f=1}^N w_{s_f,2} \mathbf{x}_f^{(s_f)}\| < \frac{\varepsilon}{3\beta}) &+ \\ \mathbb{P}(\|\mathbf{x}_0 - \sum_{s_0=1}^N w_{s_0,1} \mathbf{x}_0^{(s_0)}\| < \frac{\varepsilon}{3\beta}) &\rightarrow 1 \text{ as } N \rightarrow \infty, \end{aligned}$$

where  $\beta = \max\{B_{\mathcal{A}}, 1, \|\mathbf{x}_0\|\}$ . ■

The previous result states that if  $\mathbf{x}_f$  is sample reachable from  $\mathbf{x}_0$  in unit time,  $\mathbf{x}_f$  is also reachable under (2.1). It is worthwhile to note that we are not approximating  $\mathbf{B}$  using a stochastic ensemble approximation. If we did, then essentially we would introduce convolution-like sums which makes the problem more complicated. Further, if  $\mathbf{x}_f$  is reachable from  $\mathbf{x}_0$  only in multiple time steps, then the matter of non-unique control sequences poses another problem.

Note that the claim in Lemma 3.2 works for any sequence of random variables that form stochastic ensemble approximations as defined in Definition 2.2. In the next section we study stochastic ensemble systems that are generated from sampling the DLTI system parameters in (2.1).

#### 4 On Stochastic Ensemble Systems Generated through Sampling

Here, we deal with a particular kind of stochastic ensemble system whose realizations are produced from samples of (2.1). This corresponds to the following scenario. Consider the individual components of the states to be nodes and the  $\mathbf{A}$  matrix describing the interconnection between them. Suppose each node is aware of which neighbors are antagonistic and which neighbors are cooperative, but is unaware of the absolute magnitude of influence of its neighbors. The realizations of  $\widehat{\mathbf{A}}$  are hence obtained based on the relative order of influence of other nodes on a particular node.

In such a scenario, we assume that the stochastic ensemble system is produced in the following way. Suppose the vectors and matrices are transformed into probability mass functions over an underlying sample space. This mass function produces samples from the sample space which correspond to the realizations of the stochastic ensemble system. In particular, we first define a function  $\gamma : \mathbb{R}^n \rightarrow \mathbb{R}$ , as  $\gamma(\mathbf{w}) := \mathbf{1}^{\top} \text{abs}(\mathbf{w})$ , and with a slight abuse of notation, we consider  $\gamma : \mathbb{R}^{n \times m} \rightarrow \mathbb{R}$  as  $\gamma(\mathbf{M}) := \max_{j \in \{1, \dots, m\}} \mathbf{1}^{\top} \text{abs}(\mathbf{M}_{*j})$ . This helps us in describing a *vector sampling scheme*

procedure through the following lemma, whose proof omitted since it is trivial.

LEMMA 4.1. (*Vector sampling scheme*). Consider a vector  $\mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ . Then, let  $\bar{\mathbf{w}} := (1/\gamma(\mathbf{w}))\text{abs}(\mathbf{w})$ . Consider  $\bar{\mathbf{w}}$  as a mass function on the sample space  $\Omega = \{1, \dots, n\}$  and let  $\{s_1, \dots, s_N\}$  be  $N$  samples of  $\Omega$ . To each sample  $s_i$ , associate a vector  $\mathbf{s}^{(i)}$  as

$$[\mathbf{s}^{(i)}]_j := \begin{cases} \gamma(\mathbf{w}), & \text{if } j = s_i \text{ and } [\mathbf{w}]_j > 0; \\ -\gamma(\mathbf{w}), & \text{if } j = s_i \text{ and } [\mathbf{w}]_j < 0; \\ 0, & \text{otherwise.} \end{cases}$$

Then, the mean of the random variable corresponding to the  $i^{\text{th}}$  component of  $\sum_{i=1}^N \mathbf{s}^{(i)}$  is  $\mathbf{w}_i$ . Finally,  $\lim_{N \rightarrow \infty} \mathbb{P}(\|\mathbf{w} - \frac{1}{N} \sum_{i=1}^N \mathbf{s}^{(i)}\| > \varepsilon) \rightarrow 0$ , for each  $\varepsilon > 0$ . ■

Note that for matrices, the same procedure can be performed via a similar transformation by considering each row (or column) separately as vectors.

The stochastic ensemble system generated in this process has nice properties related to the system (2.1). First, it retains the sparsity structure of the DLTI system (2.1). Moreover, since each component is a Bernoulli random variable, using known properties we can provide convergence rates for each realization. Next, we analyze the variance of the average control computed considering stochastic ensemble approximations produced as in Lemma 4.1. The next result follows from Hoeffding's inequality.

LEMMA 4.2. (*Variance of average control*). Let  $\mathcal{X}_0, \mathcal{X}_f$  and  $\mathcal{A}$  be stochastic ensemble approximations of  $\mathbf{x}_0, \mathbf{x}_f$  and  $\mathbf{A}$  respectively generated using Lemma 4.1. Suppose  $\{\hat{\mathbf{u}}(\sigma, 0)\}_{\sigma \in \mathbb{Z}_{>0}}$  satisfy the hypothesis of Lemma 3.2 and let  $\mathbf{v}$  be as in (3.9). Then for each  $\varepsilon > 0$ ,

$$\mathbb{P}(\|\mathbf{B}(\mathbf{v} - \frac{1}{N} \sum_{\sigma=1}^N \mathbf{u}^{(s)})\| > \varepsilon) \leq 2C[\exp(-\frac{2N\varepsilon^2}{9\beta^2\gamma(\mathbf{x}_f)^2}) + \exp(-\frac{2N\varepsilon^2}{9\beta^2\gamma(\mathbf{x}_0)^2}) + \exp(-\frac{2N\varepsilon^2}{9\beta^2\gamma(\mathbf{A})^2})],$$

where  $C$  is a constant dependent only on  $n$ .

*Proof.* First we characterize the convergence rates of the realizations of  $\mathcal{X}_0, \mathcal{X}_f$ , and  $\mathcal{A}$  using the procedure in Lemma 4.1. Note that

$$(4.11) \quad \begin{aligned} \|\mathbf{x}_0 - \frac{1}{N} \sum_{s_0=1}^N \mathbf{x}_0^{(s_0)}\| &\leq \eta(n) \max_i \|\mathbf{x}_0\|_{i - \frac{1}{N} \sum_{s_0=1}^N [\mathbf{x}_0^{(s_0)}]_i}, \\ \|\mathbf{x}_f - \frac{1}{N} \sum_{s_f=1}^N \mathbf{x}_f^{(s_f)}\| &\leq \eta(n) \max_i \|\mathbf{x}_f\|_{i - \frac{1}{N} \sum_{s_f=1}^N [\mathbf{x}_f^{(s_f)}]_i}, \end{aligned}$$

where  $\eta(n)$  is a constant dependent on  $n$  such that  $\|\mathbf{w}\| \leq \eta(n)\|\mathbf{w}\|_\infty, \forall \mathbf{w} \in \mathbb{R}^n$ . Moreover,

$$(4.12) \quad \begin{aligned} \|\mathbf{A} - \frac{1}{N} \sum_{s_A(1)=1}^N \hat{\mathbf{A}}(s_A(1))\| &\leq \\ \mu(n) \max_{j \in \{1, \dots, n\}} \sum_{i \in \{1, \dots, n\}} \|\mathbf{A}\|_{ij - \frac{1}{N} \sum_{s_A(1)=1}^N [\hat{\mathbf{A}}(s_A(1))]_{ij}}, \end{aligned}$$

where  $\mu(n)$  is a constant dependent on  $n$  such that  $\|\mathbf{M}\| \leq \mu(n)\|\mathbf{M}\|_\infty, \forall \mathbf{M} \in \mathbb{R}^{n \times n}$ .

Next, by using Hoeffding's inequality [23] and exploiting the Bernoulli nature of the random variables, the right hand side of the inequalities in (4.11) and (4.12) can be bounded in probability for all  $i \in \{1, \dots, n\}$  and for each  $\varepsilon > 0$  as,

$$\begin{aligned} \mathbb{P}(\|\mathbf{x}_0\|_{i - \frac{1}{N} \sum_{s_0=1}^N [\mathbf{x}_0^{(s_0)}]_i} > \varepsilon) &\leq 2\exp(-\frac{2N\varepsilon^2}{\gamma(\mathbf{x}_0)^2}), \\ \mathbb{P}(\|\mathbf{x}_f\|_{i - \frac{1}{N} \sum_{s_f=1}^N [\mathbf{x}_f^{(s_f)}]_i} > \varepsilon) &\leq 2\exp(-\frac{2N\varepsilon^2}{\gamma(\mathbf{x}_f)^2}), \\ \mathbb{P}(\|\mathbf{A}\|_{ij - \frac{1}{N} \sum_{s_A(1)=1}^N [\hat{\mathbf{A}}(s_A(1))]_{ij}} > \varepsilon) &\leq 2\exp(-\frac{2N\varepsilon^2}{\gamma(\mathbf{A})^2}) \end{aligned}$$

Applying triangle inequality returns the results. ■

We can use the proof technique here to also characterize the rate of convergence in Lemma 3.1 considering stochastic ensemble approximations produced as in Lemma 4.1.

LEMMA 4.3. (*Variance of state trajectory*). Let  $\mathcal{X}_0, \mathcal{A}$  and  $\mathcal{A}$  be stochastic ensemble approximations of  $\mathbf{x}_0, \mathbf{A}$  and  $\mathbf{B}$  respectively, generated using Lemma 4.1. Then  $\mathbf{x}(K)$  in Lemma 3.1 satisfies, for each  $\varepsilon > 0$ ,

$$\mathbb{P}(\|\mathbf{x}(K) - \frac{1}{N} \sum_{\sigma=1}^N \hat{\mathbf{x}}(\sigma, K)\| > \varepsilon) \leq 2C(e^{c_1} + e^{c_2} + e^{c_3}),$$

where  $c_1 = -\frac{2N\varepsilon^2}{9\|\mathbf{u}\|^2\gamma(\mathbf{B})^2}$ ,  $c_2 = -\frac{2N\varepsilon^2}{9\|\mathbf{A}\|^2\gamma(\mathbf{x}_0)^2}$ ,  $c_3 = -\frac{2N\varepsilon^2}{9\|\mathbf{x}_0\|^2\gamma(\mathbf{A})^2}$  and  $C$  only depends on  $n$ .

*Proof.* The proof follows the same arguments as in the proofs of Lemmas 3.1 and 4.2. ■

#### 4.1 Sample Averaging using Least Squares Error Minimization

In the previous section, we assigned uniform weights to each realization in order to produce the stochastic ensemble approximation. For each sample run of the process, we can make this averaging method better by introducing a least square error minimization with the samples produced from the vector sampling scheme in Lemma 4.1. Since the weights associated with the optimization problem are restricted to be in the simplex  $\mathcal{S}^n$  (as per Definition 2.2), we approach this problem in two different ways. We describe each of them next and compare their performances later.

##### Accumulated Least Squares Error (ALSE) Minimization

In this approach, we take into account all the realizations of the stochastic ensemble approximations  $\mathcal{X}_0, \mathcal{X}_f$ , and  $\mathcal{A}$  all at once. Then we compute the associated averaging weights from least squares error minimization and finally use these weights for the computed control. Next, we provide a closed form solution to the least squares error minimization problem by exploiting the entries of the vectors in Lemma 4.1.

The samples produced in Lemma 4.1 have a very particular structure to them. In fact, they have a non-zero entry in one component and have zeros everywhere

else (*i.e.* their support is a singleton set). Hence, the whole problem boils down to adjusting the weights individually for the components in order to minimize the norm distance of the weighted sum of the samples and the vector. We provide this in the next result.

**LEMMA 4.4.** (*Least square error problem solution for singleton support samples*). Consider a vector  $\mathbf{w} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ . Let  $\bar{\mathbf{w}} := (1/\gamma(\mathbf{w}))\text{abs}(\mathbf{w})$  and suppose  $\{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(N)}\}$  is a set of  $N$  sample vectors produced using the sampling scheme in Lemma 4.1. Let  $\mathbf{S}$  be the matrix whose  $i^{\text{th}}$  column is  $\mathbf{s}^{(i)}$ . Then a solution to

$$(4.13) \quad \min_{\mathbf{y} \in \mathcal{S}^N} \|\mathbf{S}\mathbf{y} - \mathbf{w}\|_2^2,$$

is given by  $\mathbf{y}^*$  with  $[\mathbf{y}^*]_j = \alpha_j/n_j$ , where  $\alpha_j = |[\mathbf{w}]_k|$  with  $k$  such that  $[\mathbf{s}^{(j)}]_k \neq 0$  and  $n_j = |\{i \in \{1, \dots, N\} \mid [\mathbf{s}^{(i)}]_k \neq 0\}|$ .

*Proof.* Note that the problem is convex in  $\mathbf{y}$  and that the set  $\{1, \dots, n\}$  can be split into two disjoint sets  $\mathcal{D}_1 := \{p \in \{1, \dots, n\} \mid \exists q \text{ s.t. } [\mathbf{s}^{(q)}]_p \neq 0\}$  and  $\mathcal{D}_2 := \{1, \dots, n\} \setminus \mathcal{D}_1$ . Now, since the support of  $\mathbf{s}^{(i)}$  is a singleton, the problem can be rewritten as

$$\min_{\mathbf{y} \in \mathcal{S}^N} \sum_{i \in \mathcal{D}_1} (\sum_{j \in s(i)} [\mathbf{s}^{(j)}]_i y_j - \mathbf{w}_i)^2 + \sum_{i \in \mathcal{D}_2} \mathbf{w}_i^2 =: J(\mathbf{y}),$$

where  $s(i) := \{p \in \{1, \dots, N\} \mid [\mathbf{s}^{(p)}]_i \neq 0\}$ . It is easy to see that for the solution  $\mathbf{y}^*$  in the claim, the value of the cost function is  $J(\mathbf{y}^*) = \sum_{i \in \mathcal{D}_2} \mathbf{w}_i^2$ . As this is the minimum value of  $J(\cdot)$ , the proof is complete. ■

It is worth noting that if the set  $\mathcal{D}_2$  in the proof of Lemma 4.4 becomes empty, then  $J(\mathbf{y}^*) = 0$ . This means that if the samples are rich enough so that all the non-zero entries of the original vector appear at least once in the set  $\{\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(N)}\}$ , then the weights assigned using the least squares problem recreates the original vector perfectly. This does not translate to the case for matrices as then the associated weights to the matrix samples are coupled using multiple non-zero entries across the matrix, as illustrated in Section 5.

**Streaming Least Squares Error (SLSE) Minimization** In this case, the controls are attained from sequentially changing the realizations of the vectors and matrices. For the uniform weights case, the solution can be made better by adding more samples. For the ALSE minimization case, the weights have to be determined beforehand and then the controls computed, since we do not assign the same weights to each sample. Moreover, as the number of samples increases, the weights assigned to the previous samples need to be recomputed in order to satisfy (4.13).

However for the SLSE minimization case, the weights to the controls can be updated sequentially in order to construct a sub-optimal solution to (4.13).

Essentially, suppose  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are estimates of  $\mathbf{E}^*$ , with  $\mathbf{E}_1, \mathbf{E}_2$  and  $\mathbf{E}^*$  being comparable entities from  $\{\mathcal{X}_0, \mathcal{X}_f, \mathcal{A}\}$ , *i.e.* they are arbitrary elements in the same sequence. Then we solve for

$$(4.14) \quad \min_{(w_1, w_2) \in \mathcal{S}^2} \|w_1 \mathbf{E}_1 + w_2 \mathbf{E}_2 - \mathbf{E}^*\|_2^2.$$

The computed controls are updated using the same weights. We explain this in Algorithm 4.1.

**ALGORITHM 4.1.** Sequential control computation with least square error estimation

**Require:**  $\mathcal{X}_0, \mathcal{X}_f, \mathcal{A}, \mathbf{x}_0, \mathbf{x}_f, \mathbf{A}, \mathbf{B}, \text{errorBound}$

**Ensure:** control  $\mathbf{v}(k)$

```

1: error  $\leftarrow \infty$ ;  $\hat{\mathbf{x}}_0 \leftarrow \hat{\mathbf{x}}(\sigma, 0) \in \mathcal{X}_0$ ;
2:  $\hat{\mathbf{x}}_f \leftarrow \mathbf{x}_f(s_f(\sigma)) \in \mathcal{X}_f$ ;  $\hat{\mathbf{A}} \leftarrow \hat{\mathbf{A}}(s_A(\sigma, 1)) \in \mathcal{A}$ ;
3:  $\mathbf{v} \leftarrow \hat{\mathbf{v}}$  such that  $\hat{\mathbf{x}}_f = \hat{\mathbf{A}}\hat{\mathbf{x}}_0 + \mathbf{B}\hat{\mathbf{v}}$ 
4: while error  $>$  errorBound do
5:   Randomly choose  $i \in \{1, 2, 3\}$ 
6:   if  $i = 1$  then
7:     update  $\hat{\mathbf{x}}_0$ 
8:      $(w_1^*, w_2^*) \leftarrow$  solution of (4.14) with  $\mathbf{E}_1 = \text{old } \hat{\mathbf{x}}_0$ ,
        $\mathbf{E}_2 = \text{new } \hat{\mathbf{x}}_0$  and  $\mathbf{E}^* = \mathbf{x}_0$ .
9:   else if  $i = 2$  then
10:    update  $\hat{\mathbf{x}}_f$ 
11:     $(w_1^*, w_2^*) \leftarrow$  solution of (4.14) with  $\mathbf{E}_1 = \text{old } \hat{\mathbf{x}}_f$ ,
       $\mathbf{E}_2 = \text{new } \hat{\mathbf{x}}_f$  and  $\mathbf{E}^* = \mathbf{x}_f$ .
12:   else if  $i = 3$  then
13:    update  $\hat{\mathbf{A}}$ 
14:     $(w_1^*, w_2^*) \leftarrow$  solution of (4.14) with  $\mathbf{E}_1 = \text{old } \hat{\mathbf{A}}$ ,
       $\mathbf{E}_2 = \text{new } \hat{\mathbf{A}}$  and  $\mathbf{E}^* = \mathbf{A}$ .
15:   end if
16:    $\mathbf{v}' \leftarrow \hat{\mathbf{v}}$  such that  $\hat{\mathbf{x}}_f = \hat{\mathbf{A}}\hat{\mathbf{x}}_0 + \mathbf{B}\mathbf{v}'$ 
17:   error  $\leftarrow \| (w_1^* - 1)\mathbf{v} + w_2^* \mathbf{v}' \|$ 
18:    $\mathbf{v} \leftarrow w_1^* \mathbf{v} + w_2^* \mathbf{v}'$ 
19: end while

```

We conclude this section by comparing the averaging methods provided earlier.

**4.2 Comparison between the Averaging Methods** Since the weights obtained in the least squares error minimization are obtained from an optimization problem, we can compare their convergence rate with the uniform averaging scheme.

**LEMMA 4.5.** (*Comparison between averaging methods*). Suppose  $\{\mathbf{E}_i\}_{i=1}^N$  are estimates of  $\mathbf{E}^*$ , with  $\{\mathbf{E}_i\}_{i=1}^N$  and  $\mathbf{E}^*$  being comparable entities from  $\{\mathcal{X}_0, \mathcal{X}_f, \mathcal{A}\}$ . Let  $\mathbf{F}_{\text{uniform}} := (1/N) \sum_{i=1}^N \mathbf{E}_i$ . Let  $\mathbf{F}_{\text{ALSE}} := \sum_{i=1}^N w_i^* \mathbf{E}_i$  with  $[w_1^*, \dots, w_N^*]$  being a solution of

$$(4.15) \quad \min_{[w_1, \dots, w_N] \in \mathcal{S}^N} \|\mathbf{E}^* - \sum_{i=1}^N w_i \mathbf{E}_i\|_2^2.$$

Finally, let  $\mathbf{F}_{\text{SLSE}} := \mathbf{F}(N)$  where  $\mathbf{F}(N)$  is a solution to the difference equation  $\mathbf{F}(n) = v_1^* \mathbf{F}(n-1) + v_2^* \mathbf{F}(n-2)$  starting from  $\mathbf{F}(1) = \mathbf{E}_1, \mathbf{F}(2) = \mathbf{E}_2$  and where  $v_i^*$  comes from Algorithm 4.1. Then,

$$(4.16) \quad \|\mathbf{E}^* - \mathbf{F}_{\text{ALSE}}\| \leq \|\mathbf{E}^* - \mathbf{F}_{\text{SLSE}}\| \leq \|\mathbf{E}^* - \mathbf{F}_{\text{uniform}}\|.$$

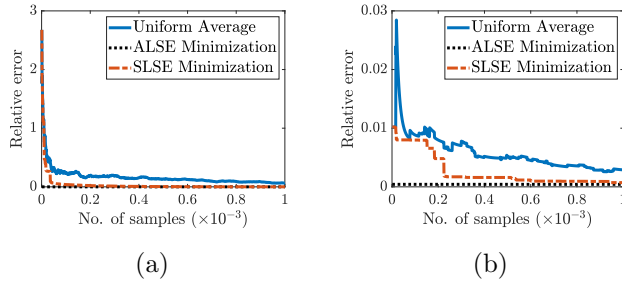


Figure 1: Relative error in estimation versus number of samples for example in Section 5.1. (a) Relative error in estimated initial condition  $\mathbf{x}_0$ . (b) Relative error in estimated system matrix  $\mathbf{A}$ .

*Proof.* Note that  $[v_1^*, \dots, v_N^*]$  is a feasible solution to (4.15). This results in the first inequality. The second inequality holds since  $(1/N)\mathbf{1}$  is a feasible solution to the optimization problem in Algorithm 4.1. ■

## 5 Simulations and Analysis

In this section, we provide simulation examples to validate our results in two scenarios. In the first case, we verify Lemma 4.5 in a power systems example. In the second case, we study the error in the computed state for a system with large dimensions. We use CVX MATLAB toolbox [24] to solve the optimization problems.

**5.1 Error in Computed Control** Here, we take an example system in the form of (2.1) from [25] using MATLAB's power toolbox. The system dimensions are  $n = 10$  and  $m = 5$  (with  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ) and we choose the initial condition randomly.

First, we verify the ordering in estimation errors for the different averaging methods in Figure 1. This is in accordance with (4.16). Moreover, note that in Figure 1(a), it is possible to obtain *zero* relative error for the initial condition estimation using ALSE Minimization and SLSE Minimization as per the discussion following Lemma 4.4. Next, we deal with the errors in the computed control and the computed state. The goal is to drive the state close to the origin in one time step. The actual control that does this was computed by taking  $\mathbf{u}_{\text{actual}} = \mathbf{B}^\dagger(-\mathbf{A}\mathbf{x}_0)$ , where  $\mathbf{B}^\dagger$  denotes the Moore-Penrose pseudoinverse. Then the three different controls  $\mathbf{u}_{\text{computed}}$  were computed using the methods listed in this paper (*i.e.* uniform average, ALSE minimization and SLSE minimization). The comparison of the relative error in computed control with respect to the number of samples is shown in Figure 2(a). Note that the  $\mathbf{u}_{\text{computed}}$  for the ALSE minimization case was computed after considering all the samples. We also computed the relative error in the computed state with respect to the number of samples and present the results in in Figure 2(b). It is worth mentioning that the error

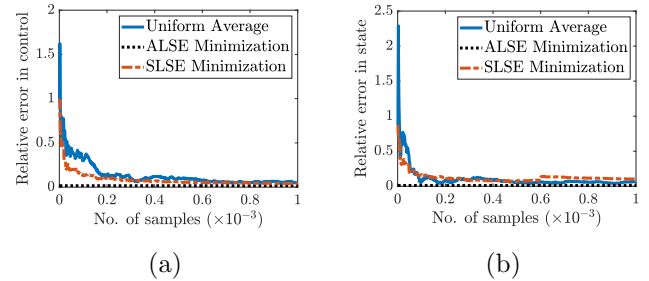


Figure 2: Relative error versus number of samples for example in Section 5.1. (a) Relative error in control  $\frac{\|\mathbf{u}_{\text{computed}} - \mathbf{u}_{\text{actual}}\|}{\|\mathbf{u}_{\text{actual}}\|}$ . (b) Relative error in state.

in the computed control and the computed states do not always follow the same ordering as the estimation errors of the initial condition and the system matrix.

**5.2 Error in Computed State** Here, we simulate our algorithms in a trajectory tracking problem for a system with large dimensions. Notice that the sample averaging method works for one time step controls only. Thus, it can be used to track a predefined trajectory of states. For this example, we take the system dimensions as  $n = 100$ , and  $m = 80$  and generate  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ,  $\mathbf{x}(0) \in \mathbb{R}^n$ ,  $\mathbf{x}(1) \in \mathbb{R}^n$ , and  $\mathbf{x}(2) \in \mathbb{R}^n$  randomly. The matrices  $\mathbf{A}$ , and  $\mathbf{B}$  are used as the system matrices and  $\mathbf{x}(0)$  is used as the initial condition. The states  $\mathbf{x}(1)$ , and  $\mathbf{x}(2)$  are used as the reference trajectory.

Similar to the previous case, for each  $t \in \{0, 1\}$ , the control for each realization  $\mathbf{A}^{(s)}$ ,  $\mathbf{x}^{(s)}(t)$ , and  $\mathbf{x}^{(s)}(t+1)$  is computed as  $\mathbf{u}_{\text{computed}} = \mathbf{B}^\dagger(\mathbf{x}^{(s)}(t+1) - \mathbf{A}^{(s)}\mathbf{x}^{(s)}(t))$ . The relative error in the computed states with respect to the number of samples is given in Figure 3. Notice that (similar to what we observed in Section 5.1) the ordering given in Lemma 4.5 that holds for the estimation errors for the initial condition and the system matrix, may not hold for the computed states as shown in Figure 3. Also, the errors in Figure 3 are much larger than the errors in Figure 2. This is because, the dimensions in this case are much larger than the previous case. Using the idea of convergence of the computed control in Lemma 3.2, we can give similar results for the convergence of computed states, omitted due to the lack of space.

## 6 Conclusion

In this paper, we studied the reachability of a new class of stochastic ensemble systems, where are not required to have a continuum of ensembles. We provided a necessary condition for reachability under multiple time steps for the limiting DLTI system using approximate reachability for the stochastic ensemble system. We also provided a sufficient condition that averages the control obtained from the sample reachability condition for the

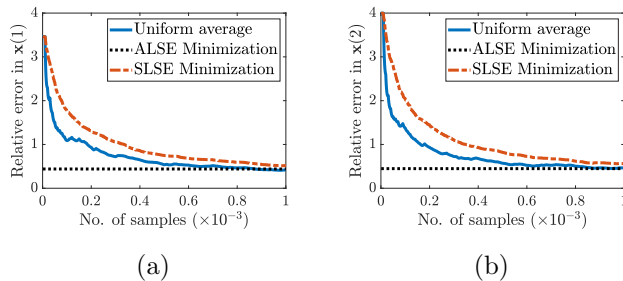


Figure 3: Relative error in state  $\frac{\|\mathbf{x}_{\text{computed}} - \mathbf{x}_{\text{actual}}\|}{\|\mathbf{x}_{\text{actual}}\|}$  versus number of samples for example in Section 5.2. (a) Relative error in  $\mathbf{x}(1)$ . (b) Relative error in  $\mathbf{x}(2)$ .

ensemble system to find a control for the DLTI system. We provided convergence rates when the stochastic ensemble systems are produced using sampling. In the future, we will extend the sufficient condition to multi-time and continuous-time cases, and will study stochastic ensemble approximations of nonlinear systems.

## References

- [1] R. Brockett, “On the control of a flock by a leader,” *Proceedings of the Steklov Institute of Mathematics*, vol. 268, no. 1, pp. 49–57, 2010.
- [2] J.-S. Li and N. Khaneja, “Control of inhomogeneous quantum ensembles,” *Physical review A*, vol. 73, no. 3, p. 030302, 2006.
- [3] S. Glaser, T. Schulte-Herbruggen, M. Sieveking, O. Schedletzky, N. Nielsen, O. Sørensen, and C. Griesinger, “Unitary control in quantum ensembles: Maximizing signal intensity in coherent spectroscopy,” *Science*, vol. 280, no. 5362, pp. 421–424, 1998.
- [4] Y. Zhai, I. Kiss, and J. Hudson, “Control of complex dynamics with time-delayed feedback in populations of chemical oscillators: Desynchronization and clustering,” *Industrial & engineering chemistry research*, vol. 47, no. 10, pp. 3502–3514, 2008.
- [5] J. S. Li and N. Khaneja, “Ensemble control of linear systems,” in *IEEE Int. Conf. on Decision and Control*, 2007, pp. 3768–3773.
- [6] K. Beauchard, J.-M. Coron, and P. Rouchon, “Controllability issues for continuous-spectrum systems and ensemble controllability of bloch equations,” *Communications in Mathematical Physics*, vol. 296, no. 2, pp. 525–557, 2010.
- [7] A. Becker and T. Bretl, “Approximate steering of a unicycle under bounded model perturbation using ensemble control,” *IEEE Transactions on Robotics and Automation*, vol. 28, no. 3, pp. 580–591, 2012.
- [8] J. Qi, A. Zlotnik, and J. S. Li, “Optimal ensemble control of stochastic time-varying linear systems,” *Systems & Control Letters*, vol. 62, no. 11, pp. 1057–1064, 2013.
- [9] A. Zlotnik and S. Li, “Synthesis of optimal ensemble controls for linear systems using the singular value decomposition,” in *2012 American Control Conference (ACC)*. IEEE, 2012, pp. 5849–5854.
- [10] S. Wang and J.-S. Li, “Fixed-endpoint minimum-energy control of bilinear ensemble systems,” in *54th Conference on Decision and Control*. IEEE, 2015, pp. 1078–1083.
- [11] M. Chertkov, V. Chernyak, and D. Deka, “Ensemble control of cycling energy loads: Markov decision approach,” *Energy Markets and Responsive Grids: Modeling, Control, and Optimization*, pp. 363–382, 2018.
- [12] A. Hassan, R. Mieth, D. Deka, and Y. Dvorkin, “Stochastic and distributionally robust load ensemble control,” *IEEE Transactions on Power Systems*, vol. 35, no. 6, pp. 4678–4688, 2020.
- [13] J. S. Li and N. Khaneja, “Ensemble control of bloch equations,” *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 528–536, 2009.
- [14] R. Brockett and N. Khaneja, “On the stochastic control of quantum ensembles,” *System theory: modeling, analysis and control*, pp. 75 – 96, 2000.
- [15] J. S. Li and J. Qi, “Ensemble control of time-invariant linear systems with linear parameter variation,” *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2808–2820, 2015.
- [16] J. S. Li, “Ensemble control of finite-dimensional time-varying linear systems,” *IEEE Transactions on Automatic Control*, vol. 56, no. 2, pp. 345–357, 2010.
- [17] L. Tie and J. S. Li, “On controllability of discrete-time linear ensemble systems with linear parameter variation,” in *American Control Conference*, 2016, pp. 6357–6362.
- [18] W. Zhang and J. S. Li, “Ensemble control on lie groups,” *SIAM J. on Control and Optimization*, vol. 59, no. 5, pp. 3805–3827, 2021.
- [19] M. Vidyasagar, “Statistical learning theory and randomized algorithms for control,” *IEEE Control Systems Magazine*, vol. 18, no. 6, pp. 69–85, 1998.
- [20] R. Tempo, G. Calafiore, and F. Dabbene, *Randomized algorithms for analysis and control of uncertain systems: with applications*. Springer, 2013, vol. 7.
- [21] M. Khajenejad and S. Yong, “Tight remainder-form decomposition functions with applications to constrained reachability and guaranteed state estimation,” *IEEE Transactions on Automatic Control*, 2023.
- [22] M. Khajenejad, F. Shoab, and S. Yong, “Guaranteed state estimation via direct polytopic set computation for nonlinear discrete-time systems,” *IEEE Control Systems Letters*, vol. 6, pp. 2060–2065, 2022.
- [23] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
- [24] M. Grant and S. Boyd, “CVX: MATLAB software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx>, Mar. 2014.
- [25] F. Pasqualetti, F. Dorfler, and F. Bullo, “Attack detection and identification in cyber-physical systems,” *IEEE Transactions on Automatic Control*, vol. 58, no. 11, p. 2715–2729, 2013.