On the External Stability of Linear Systems with Actuator Saturation Constraints, and the Decentralized Control of Communicating-Agent Networks with Security Constraints

By

ZHENG WEN

A thesis submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL ENGINEERING

WASHINGTON STATE UNIVERSITY
School of Electrical Engineering & Computer Science

AUGUST 2007

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of ZHENG WEN find it satisfactory and recommend that it be accepted.

Co-Chair

Co-Chair

Acknowledgments

I am grateful to a number of people who have made the successful culmination of my degree and this thesis possible. First and foremost, I would like to thank my two amazing advisors, Ali Saberi and Sandip Roy. Dr. Saberi has guided me in many aspects of my development as a research-oriented MS student, not only encouraging me to challenge exciting research problems and sharing his brilliant insights with me, but also helping me advance my knowledge and broaden my horizons in many areas of Systems and Control. His dedication and love for problem solving, his erudite knowledge and deep understanding of Control Theory have impressed me deeply and set a great example for me. Most of all, he has been and will always be a great mentor for me.

Since the very beginning of my study in Washington State University, I have been working very closely with Sandip: he has been a great mentor and close friend throughout my Master's work. Sandip has always spent a lot of time to help me better understand the research problems and teach me how to formulate ideas in a precise and rigorous way. Many of his cool ideas have helped me a lot to overcome the difficulties I have encountered in my research problems. After all, his enthusiasm on problem-solving, deep but clear analytical thinking have impressed and influenced me deeply.

I am very grateful to the remaining members of my thesis committee, Anton Stoorvogel and Bernie Lesieutre for their interests and encouragements on my work. I also would like to thank Electrical Engineering and Computer Science School of Washington State University for supporting me to attend the 47^{th} American Control Conference in New York City.

The research that led to this thesis could have not been accomplished without the help from my colleagues and friends. Especially, I would like to thank my current and previous fellow students in our research group–Liang Chen, Carolina Parada, James Slade, Yan Wan and Tao Yang–for their help in my study process. I am also very grateful to all my good friends in WSU. It is them that make my life in such a small town as Pullman exciting and wonderful!

Most importantly, thanks to my parents for their immense love and ever-lasting support since my childhood, and throughout my graduate study in WSU. Thanks to Liyao Wan for her support and encouraging words during the past two years.

I am also grateful for the support on my research provided by NSF grant ECS-0528882.

On the External Stability of Linear Systems with Actuator Saturation Constraints, and the Decentralized Control of Communicating-Agent Networks with Security Constraints

Abstract

by Zheng Wen, M.S.
Washington State University
August 2007

Co-Chairs: Ali Saberi and Sandip Roy

This thesis is motivated by the need to better understand the analysis and control of linear systems subject to constraints. We contribute to this field by pursuing research on two specific topics: *1)* the external stability of linear systems subject to actuator saturation, and *2)* the security of Communicating-Agent Networks (CANs). For the first topic, we study the disturbance responses of some canonical systems as the step toward better understanding of external stability of nonlinear systems. Specifically, for a saturating feedback-controlled double integrator, we delineate classes of (deterministic) disturbances that do and do not drive the state unbounded; further, for a saturating feedback-controlled single integrator driven by white noise, we fully characterize the evolution of the state's probability distribution. For the second topic, we motivate and formulate concepts of security in the context of CANs, and present some preliminary results for two classical CANs: the Single-Integrator Network and the Double-Integrator Network.

Contents

List of Figures

To my dear parents, Liyao and all my friends

Chapter 1

Introduction

Linear centralized and decentralized control systems subject to constraints are appropriate models in many areas. The constraints impacting these systems are numerous, and include the physical limitations of the plant (e.g. actuator saturation and deadzone constraints) and the requirements imposed by the designers (security constraints).

The analysis and control of linear systems with constraints is a vast field which remains incompletely understood as a whole, despite a wealth of existing works in this field (e.g. [1–5, 20]). This thesis advances the ongoing research in this field in two directions: the external stability of linear systems with actuator saturation constraints is better characterized (Chapter 2) and the concept of the security for decentralized control systems is formulated in the context of **communicating-agent networks** (Chapter 3). At last, we summarize the conclusion and suggest the future work in Chapter 4. Let us briefly introduce the two directions.

First, our study on the systems subject to actuator saturation is motivated by a need for better understanding the external stability and disturbance response of nonlinear feedback-control systems. We take some first steps in this direction by studying the external responses of some canonical systems (Saturating Static-Feedback-Controlled Single/Double Integrator) to deterministic and stochastic disturbances. Specifically, this part includes two important results: First, we have extended Shi and Saberi's work [5] on Saturating Feedback-Controlled Double Integrator by 1) finding more restrictive signal classes containing the disturbances that drive the state unbounded and 2) delineating a broad class of disturbances that do not do this. Second, we have fully characterized the time evolution of the joint probability density function (PDFs) of a Saturating Static-Feedback-Controlled Single Integrator driven by white noise for any random initial state. These results not only give insight into these two widely-used canonical systems, but also help us understand the complexity of the external stability of general nonlinear systems.

Decentralized control systems have been studied for many years (e.g. [20]), and very recently, there has been extensive research on Communicating Agent Networks (CANs) (e.g. [25, 27, 29, 30, 35])– networks of simple autonomous agents that cooperate to complete their tasks. We contend here that the security requirement is crucial in CANs and more generally in decentralized control systems. In this thesis, we carefully formulate the notions of security in the context of the CANs and develop some preliminary results for two classical CANs, the Single-Integrator Network (SIN) and the Double-Integrator Network (DIN). Specifically, we relate the notions of security for SIN/DINs with the classical notions of observability/detectability in classical linear system theory, and in turn explore how these networks' graph structures and task specifics influence the achievability of the security.

Chapter 2

On the External Stability of Linear Systems with Actuator Saturation Constraints

2.1 Background and General Problem Formulation

The interest in the study of linear systems subject to actuator saturation arises naturally from a wide recognition of inherent limitation of physical actuators. Several types of studies on the stabilization and stability of such systems have been done. These include global, semi-global internal, external and simultaneous stabilizations (e.g. [1–4]) and internal and external stability (e.g. [5]).

In this chapter, we contribute to the ongoing research on the external response of saturating static feedback-controlled systems to both deterministic and stochastic disturbances, which remain incompletely understood despite a wealth of work on the external stability of such systems. . Our effort here is not only motivated by the need for understanding the external (disturbance) responses of nonlinear and in particular saturating feedback systems for both deterministic and stochastic disturbances, but is also motivated by a continuous interest in understanding the complexity of the external stability concept and designing optimal control laws for nonlinear systems.

In this chapter, we restrict on exploring the disturbance responses of some canonical systems. Specifically, we study the deterministic disturbance response of a Saturating Static-Feedback-Controlled Double Integrator and the stochastic dynamic response of a Saturating Static-Feedback-Controlled Single Integrator driven by white noise, in Section 2.2 and Section 2.3 respectively. At last, we summarize the conclusion and suggest the future work in Section 2.4, The meaning of studying the disturbance responses of these two canonical models is twofold: first, they are classical models from the perspective of control theory and hence have been studied in detail (e.g. [5,11]) in order to better characterize the external stability of nonlinear systems; second, many realistic plants can be modeled as Saturating Double/Single Integrators, and hence understanding their external response is of great meaning in practice.

2.2 On the Deterministic Disturbance Response of a Saturating Static-Feedback-Controlled Double Integrator

2.2.1 Problem Formulation

Of particular interest in understanding the disturbance responses of saturating feedback-controlled linear systems and appropriately defining external stability for nonlinear systems, several recent articles have studied in detail the response of a canonical system, a double-integrator subject to actuator saturation that is controlled by static linear state feedback (e.g. [5, 7]), to external disturbances. Here, we continue this effort to thoroughly characterize the disturbance response of a feedback-controlled double integrator. Specifically, first, we extend the work of Shi and Saberi [5] by showing that even purely positive disturbances with arbitrarily small magnitude, as well as small disturbances with bounded time derivatives, can drive the state of the feedback system away from the origin. We then identify a broad class of sustained disturbances for which the state does remain bounded. This delineation of the disturbance signals for which the state is bounded is significant in that it allows us to evaluate precisely the external response of the canonical system to typical disturbances.

Precisely, we consider the closed-loop system

$$\Sigma: \quad \begin{cases} \dot{x_1} = x_2 \\ \dot{x_2} = \sigma(-K_1 x_1 - K_2 x_2) + \omega(t) \end{cases} \tag{2.1}$$

where the control gains K_1 and K_2 are positive constants, σ is the standard saturation function, $\omega(t)$ is a deterministic disturbance signal, and $t \in R^+$. The internal stability properties of the system Σ are well-known: the system is globally asymptotically stable (GAS). However, the work of Shi and Saberi [5] has shown that Σ is not input-to-state stable (ISS). Specifically, they have shown that for any $\delta > 0$, there exists a disturbance with $\|\omega(t)\| \leq \delta$ (henceforth denoted by $\omega(t) \in L_\infty(\delta)$) that drives the state away from the origin for some initial condition x_0.

In Subsection 2.2.2, we show that even for some more restricted classes of disturbances (e.g., small positive-valued ones), the state of the system can still be driven unbounded. In Subsection 2.2.3, we develop our main result: specifically, we identify a broad class of small disturbances for which the state does remain bounded.

2.2.2 Some Negative Results

Let us extend [5] by showing that disturbances from some typical and even more restricted classes than $L_\infty(\delta)$ can drive the state unbounded. We omit the proofs for these results because they are rather straightforward variations of the proof in [5].

We first consider disturbances that are not only bounded but have one or more derivatives that are bounded. Such derivative bounds are often appropriate in modeling real disturbances, and hence there has been some discussion on whether disturbances with limited derivatives (i.e., sluggish disturbances) have different rejection properties. In fact, as formalized below, we find that even such derivative-bounded signals can drive the state unbounded:

Theorem 1 *Consider the system* Σ *and disturbances* ω *that not only are in* $L_\infty(\delta)$ *but also satisfy* $|\frac{d^i w}{dt^i}| < \delta_i$ *for* $i = 1, \ldots, k$ *and some finite* k*. For any* $\delta > 0$, $\delta_1 > 0, \ldots, \delta_k > 0$*, there is such a disturbance that drives the state unbounded, for some initial condition* x_0*.*

Theorem 1 indicates that even when the disturbance is constrained to be small and have small derivatives, a feedback-controlled double-integrator with actuator saturation is not ISS. Briefly, we note that the theorem can be proved by smoothing out the disturbance from [5] that drives the state unbounded. The result for derivative-bounded disturbances also bears on the notion of kth-differential ISS (D^kISS) introduced recently in [8]. In a recent paper by Sontag [8], some new definitions besides ISS are given, one is differentiable k-ISS (D^kISS) stability. A system is D^kISS if there exist $\beta \in KL$ and $\gamma_i \in K_\infty$, $i = 0, 1, 2 \ldots, k$, such that the inequality

$$|x(t, x_0, u)| \leq \beta(|x_0|, t) + \sum_{i=0}^{k} \gamma_i(\|u^{(i)}\|_\infty)$$

holds for all initial values and inputs. From Theorem 1, we know the system Σ is not D^kISS, which indicates that perhaps some refinement of this definition is needed to appropriately capture external-stability notions.

Let us next consider the response of Σ to positive-valued disturbances, which are appropriate models for disturbances in e.g. industrial/chemical processes and communication networks. In some contexts, it is known that the responses of systems to strictly positive disturbances is qualitatively different from the response to general disturbances (e.g. [10]). However, as formalized next, even small positive disturbances may drive the state of the saturating double-integrator unbounded:

Theorem 2 *Consider the system* Σ *and disturbances* ω *that are in* $L_\infty(\delta)$ *and are also positive valued. For any* $\delta > 0$*, there is such a disturbance that drives the state unbounded, for some initial condition* x_0*.*

Theorem 2 can be proved simply by considering the disturbance from [5] with a constant offset added. We also note that the notions in Theorems 1 and 2 can be combined, i.e. even small positive valued disturbances with bounded derivatives can drive the state unbounded. It is worth noting that, when disturbances are constrained as in Theorems 1 and 2, the initial condition must be further from the origin than in the unconstrained case, for a disturbance to drive the system unbounded. In future work, we plan to explore further what advantages in terms of stability region and/or performance are provided by constraining disturbances in these senses.

2.2.3 A Boundedness-Preserving Signal Set

In practice, saturating double integrators subject to small external disturbances are routinely and suitably controlled using static feedback. We expect that for a very broad class of disturbances, the state of Σ will be bounded for all initial conditions. The literature contains extensive results on disturbances in L_1 and L_2 that preserve stability (specifically, ISS or some local notions of boundedness), see [6] and [9]; however, we are concerned here with persistent disturbances and so do not delve further into this literature. In this note, we aim to identify a broad class of persistent disturbances for which the system Σ has bounded state response for all initial conditions. In particular, we find that the response remains bounded when the system is impacted by small-amplitude disturbances that have bounded integral over all intervals. This is quite a broad set of small-amplitude signals, and includes periodic and quasi-periodic signals, among many others. Let us begin with a formal definition for this set of signals:

Definition 1 *For any given positive number M, the integral M-bounded signal set Ω_M is the following:*

$$\Omega_M = \{s(\cdot) \in L_\infty(1) \mid |\int_{t_1}^{t_2} s(t)dt| \leq M \ \forall\, t_2 \geq t_1 \geq 0\}.$$

We stress here that Ω_M is a parameterized set within $L_\infty(1)$, and hence for a larger M a larger set of signals is contained in Ω_M. A brief effort to identify some common types of signals that fall within the set Ω_M is worthwhile. Specifically, we can easily check that Ω_M contains 1) periodic and quasi-periodic signals with 0 mean, 2) L_1 signals, and 3) signals with spectra.

The state of the double-integrator system Σ remains bounded for all disturbances that are of sufficiently small magnitude, and that are in Ω_M. Let us formalize and prove this main result:

Theorem 3 *Consider the system Σ, for any $M > 0$, there exists a $q^*(M) > 0$ such that for any $0 < q \leq q^*$ and for any $s(t) \in \Omega_M$, $x(t; x_0, 0) \in L_\infty$ (that is, $x(t; x_0, 0)$ is bounded) for the disturbance $\omega(t) = qs(t)$ and for any initial condition x_0.*

Proof:

We re-write (2.1) using the following transformation: $y_1 = K_1x_1 + K_2x_2$, $y_2 = K_2x_2$ and $\tilde{t} = \frac{K_1}{K_2}t$. Then the system Σ is transformed to

$$\Sigma_1: \quad \begin{cases} \dot{y_1} = y_2 - \lambda[\sigma(y_1) - \omega(\tilde{t})] \\ \dot{y_2} = -\lambda[\sigma(y_1) - \omega(\tilde{t})] \end{cases} \tag{2.2}$$

where $\lambda = \frac{K_2^2}{K_1}$ is always positive, and where the derivative is taken with respect to $\tilde{t}$. Notice the integral bound is changed to $|\int_{\tilde{t_1}}^{\tilde{t_2}} s(\tilde{t})d\tilde{t}| \leq \tilde{M} = \frac{K_1}{K_2}M \ \forall\, \tilde{t_2} \geq \tilde{t_1} \geq 0$. We notice that the boundedness of $y_1(\tilde{t})$ and $y_2(\tilde{t})$ implies the boundedness of $x_1(t)$ and $x_2(t)$. Thus, to prove the theorem, let us prove the following: $y_1(\tilde{t})$ and $y_2(\tilde{t})$, which we henceforth refer to as the trajectory, are bounded for any initial condition and any disturbance $\omega(t) = qs(t)$, where $s(t) \in \Omega_M$, and $q \leq min\{0.9, \frac{0.2\lambda}{4\tilde{M}\lambda+10\tilde{M}+\frac{\lambda}{2}\tilde{M}^2}\}$. In proving this theorem, we find it convenient to consider Region $S = \{(y_1, y_2)\big| -1 \leq y_1 \leq 1, -2\lambda \leq y_2 \leq 2\lambda\}$. Also define Contour $L = \{(y_1, y_2)\big|\, y_1 = 1, y_2 > 2\lambda\} \cup \{(y_1, y_2)\big|\, y_1 = -1, y_2 < -2\lambda\}$. We note Region S is in the non-saturated region, i.e. the saturation is not activated when the state is within S. We prove the theorem in three steps:

- Step 1: First, we prove that for any initial condition $y_1(0^-), y_2(0^-)$ outside S, the trajectory will either enter S or cross L, remaining bounded in the interim.

- Step 2: Second, we prove that any trajectory that crosses L will eventually enter S, remaining bounded in the interim. Combining this result with the first step, we conclude that for any initial condition outside S, the trajectory will enter S, remaining bounded in the interim.

- Step 3: Finally, we trivially show that any trajectory starting in S will be bounded in some larger region S_1. We thus prove that the trajectory is bounded.

Step 1

We seperately consider initial conditions in 4 regions, $M_1 = \{(y_1, y_2)\big| -1 \leq y_1 \leq 1, y_2 > 2\lambda\}$, $M_2 = \{(y_1, y_2)\big| -1 \leq y_1 \leq 1, y_2 < -2\lambda\}$, $M_3 = \{(y_1, y_2)\big|\, y_1 > 1\}$, $M_4 = \{(y_1, y_2)\big|\, y_1 < -1\}$. The partition of the State Plane is illustrated in Figure 2.1.

For trajectories initially in the region M_1, we have

$$\begin{aligned} \dot{y_1} &= y_2 - \lambda y_1 + \lambda\omega \\ &\geq 2\lambda - \lambda - \lambda q = \lambda(1-q). \end{aligned} \tag{2.3}$$

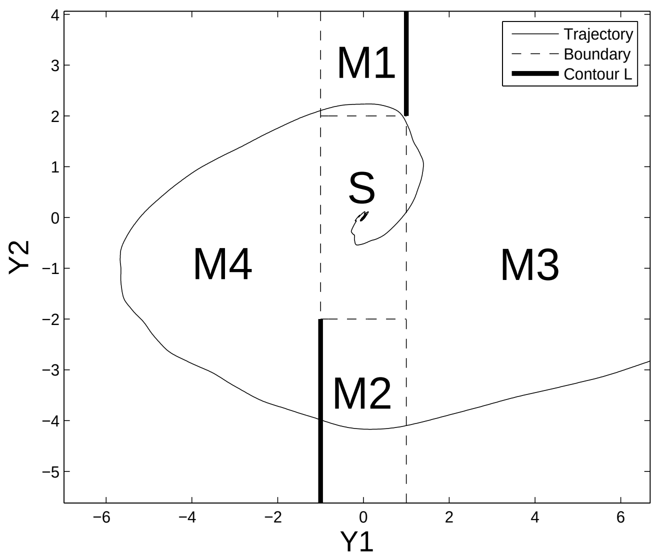


Fig. 2.1: Illustration of Partitioning the State Plane for $\lambda = 1$

Choosing $q \leq 0.9$, we have that $\dot{y_1} \geq 0.1\lambda$. Thus, any portion of a trajectory in the region M_1 satisfies $y_1(t) \geq y_1(t_0) + 0.1\lambda(t - t_0)$ for any $t > t_0$. Hence a trajectory that is initially in M_1 will either enter Region S or will definitely move rightward in M_1 and cross Contour L within a finite time of $\frac{20}{\lambda}$ time unites. Thus, the trajectory remains bounded until it enters S or crosses L, since the trajectory in non-saturated region cannot blow up in finite time. From symmetry, trajectories that are initially in the region M_2, will either enter S or cross L in the portion of $\{(y_1, y_2)\big| \, y_1 = -1, y_2 < -2\lambda\}$.

For trajectories initially in the Region M_3, again choosing $q \leq 0.9$, we have

$$\dot{y_2} = -\lambda[1 - \omega] \leq -0.1\lambda. \tag{2.4}$$

Thus y_2 will be always decreasing in region M_3 and could be positive only for finite time. When $y_2 < 0$, we have

$$\dot{y_1} = y_2 - \lambda[1 - \omega] \leq -\lambda[1 - \omega] \leq -0.1\lambda. \tag{2.5}$$

Thus it will keep decreasing and hit the boundary $y_1 = 1$. It will either enter Region S or Region M_2. Note it is impossible for the trajectory to come back to M_1 since when $y_2 > 2\lambda$, $\dot{y_1} > 0.1$ and the trajectory will definitely go rightward. As is shown above, in M_2, if the trajectory does not enter S, it will cross L. Thus the trajectory will either enter S or cross L. To show that the trajectory will be bounded in the interim, note the time needed for the trajectory to reach $y_2 = 0$ from $y_2(t_0) > 0$ is $T \leq \frac{y_2(t_0)}{0.1\lambda}$. Next when $y_2 < 0$, we have $\dot{y_1} < 0$, and we could find a bound on y_1:

$$\begin{aligned} \dot{y_1} &\leq y_2(t_0) - \lambda[1-w] < y_2(t_0) \\ y_1(t) &< y_1(t_0) + \frac{y_2^2(t_0)}{0.1\lambda}. \end{aligned} \tag{2.6}$$

Thus $y_1(t)$ is bounded in the interim. After the trajectory crosses $y_2 = 0$ inside M_3, only finite time is needed to cross $y_1 = 1$. What is more, since $-1.9\lambda \leq \dot{y_2} \leq -0.1\lambda$ in this region, y_2 is bounded and hence the trajectory cannot go unbounded in the interim. From symmetry, the similar conclusion holds for any trajectory that is initially in Region M_4. Note the only difference is that the trajectory will move upward in Region M_4 and cross the boundary $y_1 = -1$.

Based on the above proof, conclusion could be drawn that for any initial condition outside region S, the trajectory will either enter the Region S or cross the Contour L. Also notice outside S, the trajectories go in clockwise direction. This concludes the Step 1 of the proof.

Step 2

Let us consider a Point A on the Contour L (which w.l.o.g. is on the branch $\{(y_1, y_2) \big| \, y_1 = 1, y_2 > 2\lambda\}$, see Figure 2.1). From Step 1, this trajectory will go rightward from Point A and into Region M_3, and then re-cross the boundary $y_1 = 1$ at Point B. If the trajectory does not enter S during this period, it will enter Region M_2. Furthermore, if the trajectory does not enter Region S in the non-saturated region, it will cross the boundary $y_1 = -1$ (say, at Point C). Note that Point C is also on the Contour L. We shall eventually show that Point C is closer to the origin than Point A. The relative positions of the points A, B, C, D are illustrated in Figure 2.2.

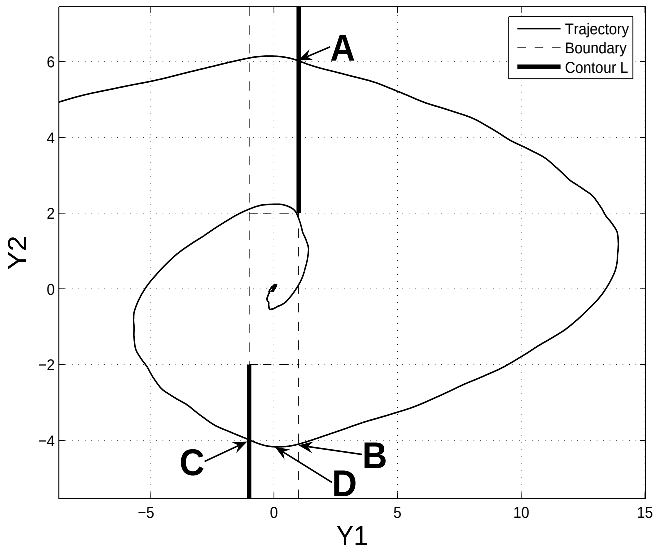


Fig. 2.2: The Relative Position of Points A, B, C, D for $\lambda = 1$

First, let us consider the Point B, and show that it is closer to the origin than A. To do so, notice that

$$\begin{aligned} y_1(t_B) - y_1(t_A) &= \int_{t_A}^{t_B} (y_2(t) - \lambda + \lambda q s(t))\, dt \\ &= 0 \end{aligned} \tag{2.7}$$

$$y_2(t_B) - y_2(t_A) = \int_{t_A}^{t_B} (-\lambda + \lambda q s(t))\, dt \tag{2.8}$$

where t_A and t_B are the times at which the trajectory is at points A and B, respectively. Actually:

$$\lambda(t_B - t_A) = \int_{t_A}^{t_B} y_2(t)dt + \lambda q \int_{t_A}^{t_B} s(t)dt. \tag{2.9}$$

Also, from (2.8) by replacing t_B to t, we have that

$$y_2(t) = y_2(t_A) + \int_{t_A}^{t} (-\lambda + \lambda q s(\tau))d\tau. \tag{2.10}$$

Combining (2.9) and (2.10), we obtain that $\lambda T =$

$$y_2(t_A)T - \frac{1}{2}\lambda T^2 + \lambda q \int_{t_A}^{t_B} \int_{t_A}^{t} s(\tau)d\tau dt + \lambda q \int_{t_A}^{t_B} s(t)dt,$$

where $T = t_B - t_A$. Using the integral bound $|\int_{t_1}^{t_2} s(t)dt| \leq \tilde{M}$, and with a little manipulation, we achieve the following inequality:

$$\frac{1}{2}\lambda T^2 + (\lambda - y_2(t_A) - \lambda q\tilde{M})T - \lambda q\tilde{M} \leq 0$$

Solving the above inequality, and using the formula $\sqrt{1+x} \leq 1 + \frac{1}{2}x \ \forall x \geq 0$, we could achieve

$$T \leq \frac{2y_2(t_A)}{\lambda} - 2 + 2q\tilde{M} + \frac{\lambda q\tilde{M}}{y_2(t_A) + \lambda q\tilde{M} - \lambda}. \tag{2.11}$$

Using (2.8) and (2.11), we can get a bound on $y_2(t_B) \geq$

$$-y_2(t_A) \quad + \quad 2\lambda - [3\tilde{M}\lambda + \frac{\lambda^2\tilde{M}}{y_2(t_A) + \lambda q\tilde{M} - \lambda}]q.$$

Since $y_2(t_A) \geq 2\lambda$, we find that

$$y_2(t_B) \geq -y_2(t_A) + 2\lambda - 4\tilde{M}\lambda q. \tag{2.12}$$

We thus notice that for $q < \frac{1}{2\tilde{M}}$, Point B is closer to the origin than Point A.

Now consider the trajectory from the Point B to the Point C. Either the trajectory enters S before hitting the boundary $y_1 = -1$, or it crosses Contour L at Point C. In the latter case, using the state-space model in the non-saturated region, we have

$$\begin{aligned} \frac{dy_2}{dy_1} &= \frac{-\lambda y_1 + \lambda\omega}{y_2 - \lambda y_1 + \lambda\omega} \\ (y_2 - \lambda y_1 + \lambda\omega)dy_2 &= (-\lambda y_1 + \lambda\omega)dy_1. \end{aligned} \tag{2.13}$$

Integrating (2.13), we have

$$\begin{aligned} & \frac{1}{2}[y_2^2(t_C) - y_2^2(t_B)] \\ = \ & \lambda[\int_{y_2(t_B)}^{y_2(t_C)} y_1 dy_2 - \int_{y_2(t_B)}^{y_2(t_C)} \omega dy_2 + \int_{y_1(t_B)}^{y_1(t_C)} \omega dy_1] \\ = \ & -\lambda^2 \int_{t_B}^{t_C} y_1^2 dt + \lambda^2 \int_{t_B}^{t_C} y_1\omega dt + \lambda \int_{t_B}^{t_C} \omega(t) \\ & [y_2(t_B) + \int_{t_B}^{t} (-\lambda y_1(\tau) + \lambda\omega(\tau))d\tau]dt. \end{aligned} \tag{2.14}$$

Notice y_1 is strictly decreasing in this course, thus the trajectory will cross $y_2 = 0$ only once. So we have $\int_{t_B}^{t_C} y_1\omega dt = \int_{t_B}^{t_D} y_1\omega dt + \int_{t_D}^{t_C} y_1\omega dt$, where D is the point where the trajectory crosses $y_2 = 0$. Note y_1 does not change sign from B to D and from D to C, thus $\int_{t_B}^{t_C} y_1\omega dt \leq 2q\tilde{M}$. Also notice that the longest possible time that is needed for the trajectory to cross Region M_2 is $\frac{20}{\lambda}$, so we have

$$
\begin{aligned}
\frac{1}{2}[y_2^2(t_C) - y_2^2(t_B)] &\leq -\lambda^2 \int_{t_B}^{t_C} y_1^2 dt + 2\lambda^2 q\tilde{M} \\
+\lambda y_2(t_B) q\tilde{M} &+ \lambda^2 q^2 \tilde{M}^2 + \lambda^2 \int_{t_B}^{t_C} \omega(t) \frac{20}{\lambda} dt \\
&\leq \lambda^2 q^2 \tilde{M}^2 + 20\lambda q\tilde{M} \\
y_2^2(t_C) \leq y_2^2(t_B) &+ 2[\lambda^2 q^2 \tilde{M}^2 + 20\lambda q\tilde{M}] \qquad (2.15)
\end{aligned}
$$

which implies (Notice $y_2(t_C)$ and $y_2(t_B)$ are negative.)

$$
\begin{aligned}
y_2(t_C) &\geq y_2(t_B) + \frac{\lambda^2 q^2 \tilde{M}^2 + 20\lambda q\tilde{M}}{y_2(t_B)} \\
&\geq y_2(t_B) - 10q\tilde{M} - \frac{\lambda}{2} q^2 \tilde{M}^2. \qquad (2.16)
\end{aligned}
$$

Thus we have

$$y_2(t_C) \geq -y_2(t_A) + 2\lambda - 4\tilde{M}\lambda q - 10q\tilde{M} - \frac{\lambda}{2} q^2 \tilde{M}^2.$$

Choosing $q \leq 0.9$ (as in Step 1), we have that

$$y_2(t_C) \geq -y_2(t_A) + 2\lambda - (4\tilde{M}\lambda + 10\tilde{M} + \frac{\lambda}{2}\tilde{M}^2)q. \qquad (2.17)$$

Let $N = 4\tilde{M}\lambda + 10\tilde{M} + \frac{\lambda}{2}\tilde{M}^2$, and further choose $q < \frac{0.2\lambda}{N}$, it follows that $y_2(t_C) \geq -y_2(t_A) + 1.8\lambda$. Thus the Point C is closer to the origin than the Point A by some non-trivial margin. From symmetry, thus we obviously see that the trajectory enters S after a finite number of clockwise circles for any given initial condition. From the proof in Step 1, the trajectory is bounded in both saturated and non-saturated regions.

Combining the proofs for Step 1 and Step 2, we have that if q satisfies the condition $q < min\{0.9, \frac{0.2\lambda}{4\tilde{M}\lambda + 10\tilde{M} + \frac{\lambda}{2}\tilde{M}^2}\}$, the trajectory will enter Region S at some time, for any initial condition outside S and for any disturbance satisfying the condition given in the theorem. This concludes the proof of Step 2.

Step 3

Consider a trajectory starting in S. We show that the trajectory remains bounded in some larger region S_1 that contains S. First, assume the trajectory goes out of S from the boundary $y_1 = 1$, say at Point A. Notice that it is impossible for the trajectory to cross the boundary $y_1 = 1$ for $y_2 \leq 0$ since $\dot{y_1} = y_2 - \lambda[1-\omega] \leq y_2 - 0.1\lambda \leq -0.1\lambda$ for $q \leq 0.9$, and so we only need consider trajectories that exit through the segment $\{y_1 = 1, y_2 > 0\}$. From (2.6), it thus follows that $y_1 \leq 1 + \frac{(2\lambda)^2}{0.1\lambda} = 1 + 40\lambda$. From (2.11), we have

$$\begin{aligned} T &\leq \frac{1}{\lambda}\Big[|y_2(t_A) - \lambda| + \lambda\tilde{M}q \\ &+ \sqrt{(|y_2(t_A) - \lambda| + \lambda\tilde{M}q)^2 + 2\lambda^2\tilde{M}q}\Big] \\ &\leq \frac{2}{\lambda}(|y_2(t_A) - \lambda| + \lambda\tilde{M}q) + \frac{\lambda\tilde{M}q}{|y_2(t_A) - \lambda| + \lambda\tilde{M}q} \\ &\leq \frac{2}{\lambda}|y_2(t_A) - \lambda| + 1.1, \end{aligned} \tag{2.18}$$

* where T is the time needed for the trajectory to re-cross the boundary $y_1 = 1$. Thus we have $T \leq 3.1$ in this case. Since we have $-1.9\lambda \leq \dot{y_2} \leq -0.1\lambda$ in M_3 (see Figure 2.1), we find $y_2 > -6\lambda$. Thus from Steps 1 and 2 of the proof, we see that the trajectory will come back to S within 2 encirclements and in finite time. Thus there exists a region N_1 in which the trajectory will be bounded. Similarly, we can prove if the trajectory crosses the boundary $y_1 = -1$, it will be bounded in a region N_2.

Then, assume the trajectory goes out of S from the upper boundary $y_2 = 2\lambda$. The trajectory either remains in M_1 until it returns to S, or crosses $y_1 = 1$ for some $y_2 > 2\lambda$, say at Point A.

*Since $q \leq \frac{0.2\lambda}{4\tilde{M}\lambda + 10\tilde{M} + \frac{\lambda}{2}\tilde{M}^2}$, we have $q \leq \frac{0.1}{2\tilde{M}}$.

From equation (2.13), we have

$$\frac{1}{2}y_2^2(t_A) - 2\lambda^2 + \frac{\lambda}{2}(1 - y_1^2(t_0)) \le \lambda^2 q^2 \tilde{M}^2 + 20\lambda q \tilde{M}$$

$$\begin{aligned} y_2(t_A) &\le 2\lambda + (10\tilde{M}q + \frac{\lambda}{2}q^2\tilde{M}^2) \\ y_2(t_A) &< 2\lambda + Nq < 2.2\lambda \end{aligned} \tag{2.19}$$

where $(y_1(t_0), 2\lambda)$ is the point where the trajectory crosses the boundary $y_2 = 2\lambda$. Notice in this case Point A is on Contour L, thus the returning point B should be closer to the origin than the Point A. Thus, from the proofs of Step 1 and Step 2, there exists some region N_3 in which this trajectory will be bounded. Similarly, a trajectory crossing the boundary $y_2 = -2\lambda$ will be bounded in some region N_4. Thus, we have shown that the trajectory is contained in a bounded region until returning to S, no matter where it escapes. Hence, the trajectory can indeed be constrained to a bounded region S_1 for all sufficiently large times; with a little effort, we find that this region is $S_1 = \{(y_1, y_2) \big| |y_1| \le 1 + 400\lambda, |y_2| \le 6.2\lambda\}$.

Combining the 3 steps in the proof, we have proven Theorem 3, for $q^* = min\{0.9, \frac{0.2\lambda}{4\tilde{M}\lambda + 10\tilde{M} + \frac{\lambda}{2}\tilde{M}^2}\}$.

□

Theorem 3 is significant in that it identifies a quite broad class of disturbances for which the response of the saturating double integrator is bounded. However, the class of permitted signals Ω_M does not include one important disturbance that leaves the state bounded: small DC disturbances leave the state bounded but are not captured in Theorem 3. In fact, it turns out that not only pure DC disturbances but also small disturbances in Ω_M with offset (i.e., with DC signal added) leave the state bounded. This more general result is stated and proved next:

Theorem 4 *Consider the system Σ. For any $M > 0$ and $d > 0$, there exists a $q^*(M, d) > 0$ such* *that, for any* $0 < q \leq q^*$ *and for any* $s(t) \in \Omega_M$*,* $x(t; x_0, 0) \in L_\infty$ *(that is,* $x(t; x_0, 0)$ *is bounded) for the disturbance* $\omega(t) = q[s(t) + d]$ *and for any initial condition* x_0*.*

Proof:

We approach the proof in 3 steps exactly as in Theorem 3, with Region S, M_1, M_2, M_3, M_4 and Contour L and Points A, B, C, D defined as before. The proofs of Step 1 and Step 3 are similar to those in Theorem 3. Further, the analysis of the trajectory from B to C in Step 2 (From Point B to Point C) is as in Theorem 3, and we need only to change the bounds on q^* appropriately. Let us thus focus on showing that Point B is strictly closer to the origin than Point A.

First, from the state-space model in the saturated region M_3, we have

$$\begin{aligned} y_1(t_B) - y_1(t_A) &= \int_{t_A}^{t_B} y_2(t) - \lambda + \lambda q[s(t) + d]dt \\ &= 0 \end{aligned} \tag{2.20}$$

$$y_2(t_B) - y_2(t_A) = \int_{t_A}^{t_B} -\lambda + \lambda q[s(t) + d]dt \tag{2.21}$$

where t_A and t_B are the times at which the trajectory is at Points A and B, respectively. From (2.20), we have

$$\begin{aligned} \lambda T &= \int_{t_A}^{t_B} y_2(t)dt + \lambda q \int_{t_A}^{t_B} s(t)dt + \lambda q d T \\ &= y_2(t_A)T - \frac{1}{2}\lambda T^2 + \lambda q \int_{t_A}^{t_B} \int_{t_A}^{t} s(\tau) d\tau dt \\ &+ \frac{\lambda}{2} q d T^2 + \lambda q \int_{t_A}^{t_B} s(t)dt + \lambda q d T \end{aligned}$$

where $T = t_B - t_A$. Using the bounds on $s(t)$ and $\int_{t_1}^{t_2} s(t)dt$, we have:

$$\frac{1}{2}(1 - qd)\lambda T^2 + (\lambda - y_2(t_A) - \lambda \tilde{M} q - \lambda d q)T - \lambda q \tilde{M}$$

≤ 0. Solving this inequality, it follows that

$$T \leq \frac{1}{(1 - qd)\lambda}[(y_2(t_A) + \lambda \tilde{M} q + \lambda d q - \lambda) + \sqrt{(y_2(t_A) + \lambda \tilde{M} q + \lambda d q - \lambda)^2 + 2\lambda^2 q \tilde{M}(1 - qd)}].$$

For $q < \frac{0.5}{|d|}$, we can always guarantee that $(1-qd)\lambda > 0$ and hence can bound T as follows:

$$\begin{aligned} T &\leq \frac{1}{(1-qd)\lambda}[2(y_2(t_A) + \lambda\tilde{M}q + \lambda|d|q - \lambda) \\ &+ \frac{\lambda^2 q\tilde{M}(1+q|d|)}{y_2(t_A) + \lambda\tilde{M}q + \lambda|d|q - \lambda}]. \end{aligned} \tag{2.22}$$

From (2.21), we can bound $y_2(t_B)$:

$$\begin{aligned} y_2(t_B) &= y_2(t_A) - \lambda T + \lambda q \int_{t_A}^{t_B} s(t)dt + \lambda q d T \\ &\geq y_2(t_A) - \lambda(1-qd)T - \lambda q\tilde{M} \end{aligned} \tag{2.23}$$

Plugging (2.22) into (2.23), we have

$$\begin{aligned} y_2(t_B) &\geq y_2(t_A) - [2(y_2(t_A) + \lambda\tilde{M}q + \lambda|d|q - \lambda) \\ &+ \frac{\lambda^2 q\tilde{M}(1+q|d|)}{y_2(t_A) + \lambda\tilde{M}q + \lambda|d|q - \lambda}] - \lambda q\tilde{M}. \end{aligned}$$

For $q < \frac{0.5}{|d|}$, we thus have $y_2(t_B) \geq$

$$\begin{aligned} -y_2(t_A) &+ 2\lambda - 2\lambda\tilde{M}q - 2\lambda|d|q - 1.5\lambda\tilde{M}q - \lambda\tilde{M}q \\ &= -y_2(t_A) + 2\lambda - q\lambda[4.5\tilde{M} + 2|d|]. \end{aligned} \tag{2.24}$$

Thus, by guaranteeing $q \leq \frac{0.2}{[4.5\tilde{M}+2|d|]\lambda}$, we can guarantee that B is closer to the origin than A by a finite margin. □

One remark about Theorem 4 is worthwhile: We can rephrase the sufficient condition in Theorem 4 so as to give it an interpretation based on time averages. To do so, we note that $s(t) \in \Omega_M$ for some finite M if and only if $|\int_{t_1}^{t_2} \omega(t)\,dt - \hat{d}(t_2 - t_1)| \leq \hat{M}$ for some $\hat{d}$ and $\hat{M}$, and for $t_2 \geq t_1$. Dividing by the length of integral, we obtain that a signal is in some Ω_M if and only if $|\frac{1}{t_2-t_1}\int_{t_1}^{t_2} \omega(t)\,dt| - \hat{d} \leq \frac{\hat{M}}{t_2-t_1}$ for some $\hat{d}$ and $\hat{M}$, for all $t_2 \geq t_1$. That is, we need the time average of the disturbance $\omega(t)$ to approach a constant at a rate that is inversely proportional to the duration over which the average is taken, for any sequence of averages of increasing duration. More simply, we can view this condition as a requirement that the time average of $\omega(t)$ reaches a limit, at a rate that is inversely proportional to the duration.

In this section, we have obtained the reassuring result the state of Σ is bounded for a broad class of disturbances. We believe that this effort constitutes a very first step toward developing useful new definitions for the external stability of nonlinear feedback systems. It is worth noting, in concluding, that the condition for boundedness presented here can trivially be formulated as another extension of ISS. However, we stress here that the important result is that we have identified a broad class of sustained external inputs for which the response is bounded; the re-formulation as an extension of ISS is not the focus for us.

2.3 On the Dynamic Response of a Saturating Static-Feedback-Controlled Single Integrator Driven by White Noise

2.3.1 Problem Formulation

In many cases, the external disturbances are modeled as a stochastic process and in particular a white noise instead of deterministic signals. Our effort in this section is motivated by the need for understanding the stochastic disturbance responses of nonlinear and in particular saturating feedback systems. In particular, several works aim to find the steady-state distribution and/or statistics of some special nonlinear stochastic systems *driven by noise* (e.g. [12] , [14]), including of feedback-controlled saturating systems (e.g. [15]). However, to the best of our knowledge, there have been no works so far that explicitly characterize the the transient response of feedback-controlled dynamic systems with actuator saturation. With the aim of better understanding the transient response, we here fully characterize the joint *probability density functions* (PDFs) of a canonical feedback-controlled saturating dynamic system, namely a single integrator with linear static feedback.

Specifically, we characterize the dynamic solution of the stochastic differential equation

$$\Sigma : \quad \begin{cases} \dot{x}(t) = \sigma[-Kx(t)] + \omega(t) \\ x(t_0) = x_0 \end{cases} , \tag{2.25}$$

where $t \in R^+$, $x(t)$ is the **state**, the **feedback gain** K is a positive real number, $\omega(t)$ is a Gaussian white noise process with zero mean and autocorrelation $R(t, \tau) = \delta(t - \tau)$, the initial state x_0 is a random variable that is independent of $\omega(t)$, and $\sigma(\cdot)$ is the standard saturation function, i.e.

$$\sigma(x) = \begin{cases} 1 & x > 1 \\ x & -1 \leq x \leq 1 \\ -1 & x < -1 \end{cases} .$$

Notice that the differential equation Σ exactly represents the dynamics of a saturating static-feedback-controlled single integrator with additive white Gaussian noise (AWGN) input $\omega(t)$. The steady-state probability distribution of $x(t)$ has been found by Fuller ([12]). In this paper, our aim is to fully characterize the state $x(t)$, which is a stochastic process, for any initial random state x_0. In other words, we will find the joint probability density function (PDF) $W_{x(t)}(x_1, t_1, x_2, t_2, \cdots, x_n, t_n)$ of the random variables $x(t_1), x(t_2), \cdots, x(t_n)$ for all $t_0 \leq t_1 < t_2 < \cdots < t_n$ (w.l.o.g.), and for every natural number n.

Our methodology for finding the joint PDFs is the following: we first find the conditional first-order PDF for $x(t)$. The first-order PDF of the state of a nonlinear stochastic system, such as Σ, is known to evolve according to a partial differential equation called the Fokker-Planck equation ([13]). For this particular system, we take the common approach of transforming the Fokker-Planck operator to a Hermitian Schrödinger operator, which makes finding the solution easier by permitting expression of the PDF in an orthogonal basis of eigenfunctions. Then we solve for and normalize the eigenfunctions of the resulting Schrödinger operator and hence derive the conditional PDF $P(x,t|x',t_0)$, $x,x' \in \Re$, $t \geq t_0$ using the eigenfunction expansion. Using the fact that $x(t)$ is Markovian†, we can thus obtain any $n^{th}-$order PDF of the stochastic process $x(t)$ for arbitrary random initial state $x(t_0)$. Thus, we fully characterize the stochastic process $x(t)$.

Similar models have also been pursued in thermodynamics and quantum mechanics, where the motions of molecules and quantum behavior of fundamental particles are described by nonlinear dynamic systems driven by stochastic processes. From the perspective of thermodynamics, Σ is the Langevin equation for Brownian motion in an environment where particles are subject to a saturating linear damping force. In this context, the joint PDF of the stochastic process $x(t)$ identifies the distribution of the particle velocities at any combination of time points. As will be shown later, the *potential* of the derived Fokker-Planck equation is parabolic in the center and linear in the tails, which is a variant on the widely-considered parabolic potential (e.g. [16]) and V-shaped potential (e.g. [17]). From the perspective of quantum mechanics, this paper solves a Schrödinger equation with a discontinuous finite parabolic well potential, see [18] for eigenanalysis of a somewhat similar potential.

The remainder of the note is as follows: in Subsection 2.3.2, we present our main results; in Subsection 2.3.3, we prove these results; in Section 2.3.4, we illustrate our results with two examples.

2.3.2 Main Result: Dynamic PDF of the State

We have three aims in presenting the results in this subsection. First, we clarify that the conditional PDF $P(x,t|x',t_0)$ of $x(t)$ given $x_0 = x'$, $t \geq t_0$ can be obtained through a eigenfunction (spectral) expansion, in particular based on transformation of the Fokker-Planck operator

†Since Σ is causal and $\omega(t)$ is white noise.

to a (Hermitian) Schrödinger operator. Second, we give a systematic procedure for finding and normalizing the eigenfunctions needed for the eigenfunction expansion. At last, we show that $W_{x(t)}(x_1, t_1, x_2, t_2, \cdots, x_n, t_n)$ can be easily obtained given conditional PDFs and initial distribution. The following theorems and lemmas summarize the main results: ‡

Theorem 5 *Consider the solution $x(t)$ to the differential equation (2.25). The conditional PDF for $x(t)$ given $x(t_0) = x'$, $t \geq t_0$ is given by*

$$\begin{aligned} P(x,t|x',t_0) &= e^{U(x')-U(x)} \sum_{i=0}^{N} \Psi_D(x';i)\Psi_D(x;i)e^{-\lambda_i(t-t_0)} \\ &+ e^{U(x')-U(x)} \frac{1}{\pi} \int_0^\infty [\Psi_E(x';u)\Psi_E(x;u) + \Psi_O(x';u)\Psi_O(x;u)] e^{-\frac{1+u^2}{2}(t-t_0)} du, \end{aligned} \tag{2.26}$$

where $N \geq 0$,

$$U(x) = \begin{cases} x - \frac{1}{2K} & x > \frac{1}{K} \\ -x - \frac{1}{2K} & x < -\frac{1}{K} \\ \frac{K}{2}x^2 & -\frac{1}{K} \leq x \leq \frac{1}{K} \end{cases},$$

and, the normalized eigenfunctions $\Psi = \Psi_D(x;i)$, $\Psi = \Psi_E(x;u)$ and $\Psi = \Psi_O(x;u)$ are solutions of the equation

$$\frac{d^2\Psi}{dx^2} + \{-\sigma[Kx]^2 + K1^+(\frac{1}{K} - x)1^+(\frac{1}{K} + x) + 2\lambda\}\Psi = 0, \tag{2.27}$$

for some $\lambda \geq 0$

1) *Specifically, the discrete eigenfunctions $\Psi_D(x;i)$ are the N solutions to (2.27) for $0 \leq \lambda < \frac{1}{2}$ that have continuous value and derivative at each x, are everywhere bounded, and are normalized so that $\int_{-\infty}^{\infty} \Psi_D^2(x;i)dx = 1$ (i.e. $\Psi_D(x;i) \in L_2(1)$).*

‡The spectral approach to solving the dynamic Fokker-Planck equation is well-known in the physics community (see e.g. [13]), however, we overview this methodology for the control/systems community. To the best of our knowledge, the computation of the eigenfunctions in this case, and the connection between saturating control systems and the dynamic Fokker-Planck equation, are novel.

2) The continuous even and odd eigenfunctions $\Psi_E(x;u)$ *and* $\Psi_O(x;u)$ *are the solutions to (2.27) for* $\lambda = \frac{1}{2}(1+u^2)$ *that have continuous value and derivative at each* x*, that are even and odd respectively, and that are normalized so* $max_{x\in(-\infty,-\frac{1}{K})\cup(\frac{1}{K},\infty)}\Psi_E(x;u) = 1$ *and* $max_{x\in(-\infty,-\frac{1}{K})\cup(\frac{1}{K},\infty)}\Psi_O(x;u) = 1$*. We note that these continuous eigenfunctions are persistent and hence are not absolutely integrable.*

In the following, we describe how the functions of $\Psi_D(x;i)$, $\Psi_E(x;u)$ and $\Psi_O(x;u)$ can be computed. Specifically, we first give two preliminary lemmas on solving (2.27) and then a theorem for finding the bases.

Lemma 1 *For any* $\lambda \geq 0$*, the solution of (2.27) for* $x \in (-\infty, -\frac{1}{K}) \cup (\frac{1}{K}, \infty)$ *is*

$$\Psi(x) = \begin{cases} C_1 e^{\sqrt{1-2\lambda}x} + C_2 e^{-\sqrt{1-2\lambda}x} & \text{for } 0 \leq \lambda < \frac{1}{2} \\ C_1 sin(\sqrt{2\lambda-1}x) + C_2 cos(\sqrt{2\lambda-1}x) & \text{for } \lambda > \frac{1}{2} \\ C_1 + C_2 x & \text{for } \lambda = \frac{1}{2} \end{cases},$$

where C_1 *and* C_2 *are coefficients to be determined from the boundary condition (continuous value and derivative at* $x = \pm\frac{1}{K}$*) and from normalization.*

Lemma 2 *The solution of (2.27) for* $x \in [-\frac{1}{K}, \frac{1}{K}]$ *is*

$$\Psi(x) = (1,0)\Phi(x,0)(D_1, D_2)^T, \tag{2.28}$$

where $\Phi(x,0) = \sum_{i=0}^{\infty} M_i(x)$*,* D_1 *and* D_2 *are coefficients to be decided from the boundary condition and normalization, and* $M_i(x)$*,* $i = 0, 1, 2, \cdots$ *are a sequence of* 2×2 *matrices that can be generated as follows:*

1) $M_0(x) = I_2$*;*

2) For even j*,* $M_{j+1}(x)$ *can be generated from* $M_j(x)$ *as follows:* $M_{j+1}(x) =$

$$\begin{bmatrix} 0 & \sum_{i=0}^{N_2(j)} \frac{q_i(j)}{i+1} x^{i+1} \\ \sum_{i=0}^{N_1(j)} p_i(j)(\frac{b}{i+1}x^{i+1} + \frac{a}{i+3}x^{i+3}) & 0 \end{bmatrix},$$

where $p_i(j)$ and $q_i(j)$ are coefficients of the polynomial entries of $M_j(x)$, specifically

$$M_j(x) = \begin{bmatrix} \sum_{i=0}^{N_1(j)} p_i(j)x^i & 0 \\ 0 & \sum_{i=0}^{N_2(j)} q_i(j)x^i \end{bmatrix},$$

3) For odd j, $M_{j+1}(x)$ can be generated from $M_j(x)$ as follows: $M_{j+1}(x) =$

$$\begin{bmatrix} \sum_{i=0}^{N_2(j)} \frac{q_i(j)}{i+1} x^{i+1} & 0 \\ 0 & \sum_{i=0}^{N_1(j)} p_i(j)(\frac{b}{i+1}x^{i+1} + \frac{a}{i+3}x^{i+3}) \end{bmatrix},$$

where $p_i(j)$ and $q_i(j)$ are coefficients of the polynomial entries of $M_j(x)$, specifically

$$M_j(x) = \begin{bmatrix} 0 & \sum_{i=0}^{N_1(j)} p_i(j)x^i \\ \sum_{i=0}^{N_2(j)} q_i(j)x^i & 0 \end{bmatrix},$$

where $a = K^2$, $b = -(2\lambda + K)$.

Lemma 3 *For any $\lambda \geq 0$, Equation (2.27) has both an even solution and an odd solution, which can be determined by selecting $(D_1, D_2) = (1, 0)$ and $(D_1, D_2) = (0, 1)$ respectively in (2.28), and choosing C_1 and C_2 in (2.28) in such a way that both $\Psi(x)$ and its derivative are continuous at $x = \pm\frac{1}{K}$ (i.e. $\Psi(x) \in C^1$).*

Theorem 6 *$\Psi_D(x; i)$, $\Psi_E(x; u)$ and $\Psi_O(x; u)$ are obtained as follows:*

1) There are a finite set of λ in $[0, \frac{1}{2})$ (say $0 \leq \lambda_0 < \lambda_1 < \cdots \lambda_N < \frac{1}{2}$) such that the coefficient C_1 of the tail solution (2.28) computed in Lemma 3 is 0 (i.e. the solution does not contain an exponentially growing term) for either the even solution or odd solution. then $\Psi_D(x; i)$ equals $C\Psi(x)$, where $\Psi(x)$ is the (non-growing) solution to (2.27) associated with λ_i and C is chosen so that $\int_{-\infty}^{\infty} \Psi^2(x)dx = 1$.

2) For any $\lambda > \frac{1}{2}$, let $u = \sqrt{2\lambda - 1}$. Normalize the even and odd solutions of (2.27) associated with λ such that their amplitudes on $(-\infty, -\frac{1}{K}) \cup (\frac{1}{K}, \infty)$ is 1. Label these normalized solutions as $\Psi_E(x; u)$ and $\Psi_O(x; u)$, respectively.

Before proving Theorem 5 and 6 and the associated lemmas, it is worthwhile to notice that –based on the classical analysis of Markov processes– we can straightforwardly compute first- and higher-order PDFs of $x(t)$ given the distribution of the initial state x_0. The following remarks formalize these notions:

Remark 1 *The first-order PDF for $x(t)$ given initial distribution $W(x, t_0) = \delta(x - x')$ (i.e. deterministic initial state $x(t_0) = x_0 = x'$) is $W(x, t) = P(x, t|x', t_0)$.*

Remark 2 *For any initial distribution $W(x, t_0)$, according to the Kolmogorov equation, the first-order PDF $W(x, t)$ is given as*

$$W(x, t) = \int_{-\infty}^{\infty} P(x, t|x', t_0)W(x', t_0)dx'.$$

Remark 3 *Using the Markov property, the n^{th}-order PDF for $x(t)$ at $t_1, t_2, \cdots, t_n$ can be found*

$$W_{x(t)}(x_1, t_1, x_2, t_2, \cdots, x_n, t_n) = \prod_{i=2}^{n} P(x_i, t_i|x_{i-1}, t_{i-1})W(x_1, t_1).$$

Thus, statistics for the state such as the autocorrelation can also be found.

Remark 4 *It is straightforwardly to check that, as $t \to \infty$, the first-order PDF $W(x, t)$ that we have found converges to the steady-state PDF reported by Fuller, namely*

$$W_S(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{1}{K}} e^{-Kx^2} & |x| \leq \frac{1}{K} \\ \frac{1}{\beta} e^{-2|x|} & |x| > \frac{1}{K} \end{cases},$$

where

$$\beta = 2[\int_0^{\frac{1}{K}} e^{-\frac{1}{K}} e^{-Kx^2} dx + \int_{\frac{1}{K}}^{\infty} e^{-2x} dx].$$

2.3.3 Proof of the Dynamic Response

In this section, we prove Theorem 5 and 6, and Lemmas 1-3. First, we review the derivation of the Fokker-Planck equation from Σ, transform the Fokker-Planck operator into a Hermitian Schrödinger operator and obtain the form of the solution based on an eigenfunction expansion of the transformed Hermitian Schrödinger operator. Then, we find and normalize the eigenfunctions of the Schrödinger operator associated with each eigenvalue $\lambda \geq 0$ and hence derive the conditional PDF $P(x, t|x_0, t_0)$.

A: Derivation of Fokker-Planck Equation, Transformation to Schrödinger Equation and Eigenfunction Expansion:

Our goal is to obtain the first-order PDF $W(x, t)$ of the state $x(t)$ of Σ at each time t. As is well-known in [13], this PDF can be shown using local balance arguments to evolve according the following Fokker-Planck equation:

$$\begin{aligned} \frac{\partial}{\partial t}W(x,t) &= \frac{\partial}{\partial x}\Big(\sigma[Kx(t)]W(x,t)\Big) + \frac{1}{2}\frac{\partial^2}{\partial x^2}W(x,t) \\ &= L_{FP}W(x,t), \end{aligned} \tag{2.29}$$

where $\mathbf{L_{FP}}$ is called the **Fokker-Planck operator**. As is classical, let us consider a solution based on variable separation, i.e. one of the form

$$W(x,t) = \sum_{\lambda\in\Lambda} b(\lambda)\varphi(x;\lambda)e^{-\lambda(t-t_0)}, \tag{2.30}$$

§ where the **eigenfunctions** $\varphi(x; \lambda)$, $\lambda \in \Lambda$ form a complete basis for real scalar functions with unit 1 norm, $\lambda \in \Lambda$ are known as **eigenvalues**, the **weights** $b(\lambda)$ are formed through the projection of $W(x, t_0)$ on the basis, and the summation in (2.30) is generalized (i.e. may include integration over a continuous set of λ). Substituting (2.30) into (2.29) leads to the time-invariant eigenvalue

§Here we assume each eigenvalue only has 1 linearly independent eigenfunction and the eigenvalues are discrete.

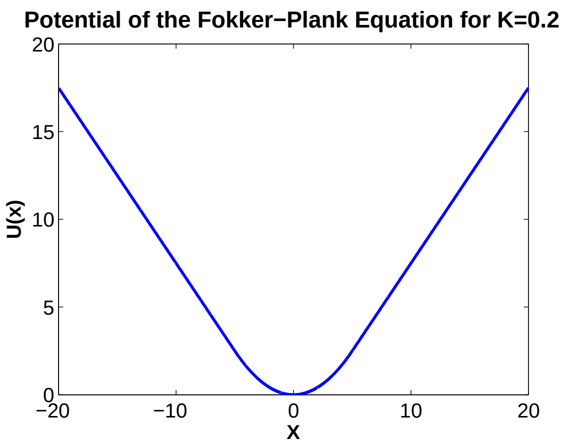


Fig. 2.3: Potential Function of the Derived Fokker-Planck Equation

equation $L_{FP}\varphi(x;\lambda) = -\lambda\varphi(x;\lambda)$. Unfortunately, the fact that $\mathbf{L_{FP}}$ is in general not a Hermitian operator makes it very difficult to compute the projection $b(\lambda)$ on each eigenfunction even if the eigenfunctions can be found analytically. Using the transformation $L = e^{U(x)}L_{FP}e^{-U(x)}$, where

$$\begin{aligned} U(x) &= -\int_0^x \sigma[-Kz]dz \\ &= \begin{cases} x - \frac{1}{2K} & x > \frac{1}{K} \\ -x - \frac{1}{2K} & x < -\frac{1}{K} \\ \frac{K}{2}x^2 & -\frac{1}{K} \leq x \leq \frac{1}{K} \end{cases}, \end{aligned} \tag{2.31}$$

is the **potential** of the Fokker-Planck equation (see Figure 2.3.3), the Fokker-Planck operator $\mathbf{L_{FP}}$ is transformed to the **Schrödinger operator L**, which is Hermitian and thus has orthogonal eigenfunctions for distinct eigenvalues. Any $\lambda \in \Lambda$ is also an eigenvalue of **L** and its associated eigenfunction is $\Psi(x;\lambda) = e^{U(x)}\varphi(x;\lambda)$. Thus, $W(x,t)$ can be expressed as

$$W(x,t) = e^{-U(x)} \sum_{\lambda \in \Lambda} B(\lambda)\Psi(x;\lambda)e^{-\lambda(t-t_0)}, \tag{2.32}$$

where $B(\lambda)$ is the projection of $e^{U(x)}W(x,t_0)$ on the eigenspace spanned by $\Psi(x;\lambda)$. We note that $B(\lambda)$ can be easily computed upon proper normalization of $\Psi(x;\lambda)$. With a little manipulation, we find the time-invariant Schrödinger equation $\mathbf{L}\Psi(x;\lambda) = -\lambda\Psi(x;\lambda)$ is equivalent to

$$\frac{d^2\Psi}{dx^2} + \{-\sigma[Kx]^2 + K1^+(\frac{1}{K} - x)1^+(\frac{1}{K} + x) + 2\lambda\}\Psi = 0, \tag{2.33}$$

where $1^+(\cdot)$ is the unit step function and since $\mathbf{L}$ is Hermitian, we need only consider $\lambda \geq 0$. Notice the above equation is same as (2.27). Since (2.33) is a second-order differential equation, any eigenvalue $\lambda \geq 0$ will have 2 linearly independent eigenfunctions, which can be orthogonalized. Noting in general that the eigenvalue set Λ required to span the function space contains both a countable set of eigenvalues Λ_D and a continuous set Λ_C with each eigenvalue being repeated at most twice, we can write $W(x,t)$ more precisely

$$W(x,t) = e^{-U(x)}\left(\sum_{\lambda_i \in \Lambda_D} B(i)\Psi(x;i)e^{-\lambda_i(t-t_0)} + \int_{\lambda\in\Lambda_C} B(\lambda)\Psi(x;\lambda)e^{-\lambda(t-t_0)}d\lambda\right). \tag{2.34}$$

From this observation, we can normalize the eigenfunctions so that

$$\int_{-\infty}^{\infty} \Psi(x;m)\Psi(x;n)dx = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \tag{2.35}$$

for m, n such that $\lambda_m, \lambda_n \in \Lambda_D$,

$$\int_{-\infty}^{\infty} \Psi(x;\lambda)\Psi(x;\lambda') = \delta(\lambda - \lambda') \tag{2.36}$$

for $\lambda, \lambda' \in \Lambda_C$,

$$\int_{-\infty}^{\infty} \Psi(x;\lambda)\Psi(x;n)dx = 0 \tag{2.37}$$

for $\lambda \in \Lambda_C$ and $\lambda_n \in \Lambda_D$. Using this normalization, we can compute the weights $B(i)$ and $B(\lambda)$ in the expansion for $W(x,t)$. To simplify this projection a bit further, let us limit ourselves to the case where the initial PDF for $x(t)$ is $\delta(x-x')$; in this case, we obtain $W(x,t) = P(x,t|x',t_0)$, from which the first-order PDF for arbitrary initial state distribution can be found using the Kolmogorov equation (2.29). Specifically, for $\lambda_i \in \Lambda_D$, we have

$$\begin{aligned} B(i) &= \int_{-\infty}^{\infty} e^{U(x)}P(x,t_0|x',t_0)\Psi(x;i)dx \\ &= \int_{-\infty}^{\infty} e^{U(x)}\delta(x-x')\Psi(x;i)dx \\ &= e^{U(x')}\Psi(x';i), \end{aligned} \tag{2.38}$$

and for $\lambda \in \Lambda_C$, we have

$$\begin{aligned} B(\lambda) &= \int_{-\infty}^{\infty} e^{U(x)} P(x, t_0 | x', t_0) \Psi(x; \lambda) dx \\ &= \int_{-\infty}^{\infty} e^{U(x)} \delta(x - x') \Psi(x; \lambda) dx \\ &= e^{U(x')} \Psi(x'; \lambda). \end{aligned} \tag{2.39}$$

We will find and normalize the eigenfunctions of the Schrödinger operator $\mathbf{L}$ in Subsection 2.3.3 by solving (2.33). The conditional PDF $P(x, t|x', t_0)$ of Σ will be found based on the general principles described above and the specific form of the eigenfunctions.

B: Finding and Normalizing the Eigenfunctions of the Schrödinger Operator

In this subsection, we derive and normalize the eigenfunctions of the Schrödinger operator $\mathbf{L}$ by solving the differential equation (2.33). First, we find the solution of (2.33) for $|x| > \frac{1}{K}$ and for $|x| \leq \frac{1}{K}$ separately. Second, the entire eigenfunction $\Psi(x)$, $x \in \Re$, is obtained by matching the solutions in the regions at the boundaries to within a normalization factor. That is, the coefficients of the solution are set so that $\Psi(x)$ has continuous value and derivative at $x = \pm\frac{1}{K}$. Third, we normalize the eigenfunctions according to Equation (2.35) and (2.36). At last, from projection coefficient formulas (2.38), (2.39) and the synthesis formula (2.32), we can easily obtain the conditional PDF (2.26).

B1: Solving the Schrödinger Equation for $|x| > \frac{1}{K}$

We now solve the Schrödinger equation (2.33) for $|x| > \frac{1}{K}$ and hence verify Lemma 1. In the two *tail regions* $x \in (-\infty, -\frac{1}{K})$ and $x \in (\frac{1}{K}, \infty)$, Equation (2.33) becomes:

$$\frac{d^2\Psi}{dx^2} + (-1 + 2\lambda)\Psi = 0. \tag{2.40}$$

Trivially, the solution to the linear constant-coefficient differential equation (2.40) is

$$\Psi(x) = \begin{cases} C_1 e^{\sqrt{1-2\lambda}x} + C_2 e^{-\sqrt{1-2\lambda}x} & \text{for } 0 \le \lambda < \frac{1}{2} \\ C_1 sin(\sqrt{2\lambda - 1}x) + C_2 cos(\sqrt{2\lambda - 1}x) & \text{for } \lambda > \frac{1}{2} \\ C_1 + C_2 x & \text{for } \lambda = \frac{1}{2} \end{cases}, \tag{2.41}$$

where C_1 and C_2 are coefficients to be determined from the boundary condition at $x = -\frac{1}{K}$ and $x = \frac{1}{K}$. Thus we complete the proof of Lemma 1.

B2: Solving the Schrödinger Equation for $|x| \le \frac{1}{K}$ by the Peano-Baker Series

We now solve the Schrödinger equation (2.33) for $|x| \le \frac{1}{K}$, and hence prove Lemma 2. For the *central region* $x \in (-\frac{1}{K}, \frac{1}{K})$, Equation (2.33) becomes:

$$\frac{d^2\Psi}{dx^2} + \{-(Kx)^2 + K + 2\lambda\}\Psi = 0. \tag{2.42}$$

Notice the Equation (2.42) is a non-constant-coefficient linear differential equation. Letting $y_1(x) = \Psi(x)$, $y_2(x) = \frac{d\Psi}{dx}(x)$, $a = K^2$, $b = -(2\lambda + K)$, $y = (y_1, y_2)^T$ and

$$A(x) = \begin{pmatrix} 0 & 1 \\ ax^2 + b & 0 \end{pmatrix},$$

we can rewrite the differential equation (2.42) in state-space form:

$$\dot{y} = A(x)y. \tag{2.43}$$

The fact that $A(x_1)$ and $A(x_2)$ are not commutative makes it difficult to obtain a closed-form solution to Equation (2.43). However, from classical linear system theory, we can find the solution as $y(x) = \Phi(x, 0)y(0)$, where the position transition matrix $\Phi(x, 0)$ can be found using the Peano-Baker Series [19]. That is, defining $M_0(x) = I$ and $M_j(x) = \int_0^x A(s)M_{j-1}(s)ds$, we have $\Phi(x, 0) = \sum_{i=0}^{\infty} M_i(x)$. For $A(x)$ defined in (2.43), we thus obtain the algorithm described in Lemma 2 with just a little algebra. Letting $y(0) = (D_1, D_2)^T$ and noting that $\Psi(x) = y_1(x)$, we obtain (2.28).

Remark 5 If $(D_1, D_2) = (1, 0)$, the resultant $y(x)$ is even on $[-\frac{1}{K}, \frac{1}{K}]$, while if $(D_1, D_2) = (0, 1)$, the resultant $y(x)$ is odd on $[-\frac{1}{K}, \frac{1}{K}]$.

B3: Deriving and Normalizing the Full Eigenfunctions

So far, we have obtained the solutions for the transformed Schödinger equation (2.33) in three separate regions: $(-\infty, -\frac{1}{K})$,$[-\frac{1}{K}, \frac{1}{K}]$ and $(\frac{1}{K}, \infty)$, for each eigenvalue λ. The corresponding eigenfunction $\Psi(x;\lambda)$ for eigenvalue λ can be obtained by matching the solution at boundaries, i.e. by ensuring that the eigenfunction and its derivative are continuous at the boundaries $x = \pm\frac{1}{K}$ and hence at all points ($\Psi(x;\lambda) \in C^1(-\infty,\infty)$). Since the Schödinger operator is second-order, each eigenvalue $\lambda \geq 0$ is repeated twice, and so has an even and an odd eigenfunction $\Psi_E(x;\lambda)$ and $\Psi_O(x;\lambda)$ associated with it. These 2 eigenfunctions per eigenvalue together form an orthogonal base for the eigenspace associated with λ.

From the classical observation that the bounded solutions to (2.33) form a complete basis, the transformed initial condition $e^{U(x')}\delta(x-x')$ can be written in terms of the exponentially-decreasing eigenfunctions for $\lambda < \frac{1}{2}$ and the sinusoid-tailed eigenfunctions for $\lambda \geq \frac{1}{2}$. As discussed next, the bounded (exponentially-decreasing) eigenfunctions for $\lambda < \frac{1}{2}$ form a finite set, while there are a continuum of eigenfunctions for $\lambda \geq \frac{1}{2}$.

Specifically, for $\lambda < \frac{1}{2}$, each λ can have at most one linearly independent exponentially decreasing eigenfunction, which is either the even eigenfunction $\Psi_E(x;\lambda)$ or the odd eigenfunction $\Psi_O(x;\lambda)$. By considering the form of the eigenfunction of either side of the boundary, we can check (with a little effort) that the total number of eigenvalues that have an exponentially decreasing eigenfunction will be finite, say N. Thus, we can label these eigenvalues as $\lambda_0, \lambda_1, \lambda_2, \cdots, \lambda_{N-1}$. For any integer $i = 0, 1, 2, \cdots, N-1$, we label the exponentially decreasing eigenfunction associated with λ_i as $\Psi_D(x;i)$. According to (2.35), we need to normalize $\Psi_D(x;i)$ such that $\int_{-\infty}^{\infty} \Psi_D^2(x;i)dx = 1$ (i.e. $\Psi_D(x;i) \in L_2(1)$). Also notice $\lambda_0 = 0$ always has an exponentially decreasing eigenfunction equal to the steady-state distribution by a scaling factor for any $K > 0$.

For $\lambda \geq \frac{1}{2}$, each λ has both an even and an odd eigenfunction that is bounded but persistent, and hence there is a continuum of eigenfunctions for $\lambda \geq \frac{1}{2}$. For convenience, we use the change of variables $u = \sqrt{2\lambda - 1}$ in the basis decomposition. We thus use $\Psi_E(x; u)$ and $\Psi_O(x; u)$ to denote the even and odd eigenfunctions associated with eigenvalue $\lambda = \frac{1}{2}(1 + u^2)$. Substituting into the normalization equations and using the standard integration of sinusoid products (as in Fourier theory), we obtained that the eigenfunctions should be normalized so that the amplitude of the sinusoid oscillation for $|x| > \frac{1}{K}$ is 1:

$$\begin{aligned}
\frac{1}{\pi}\int_{-\infty}^{\infty} \Psi_E(x; u)\Psi_E(x; u')dx &= \delta(u - u') \\
\frac{1}{\pi}\int_{-\infty}^{\infty} \Psi_O(x; u)\Psi_O(x; u')dx &= \delta(u - u'),
\end{aligned}$$

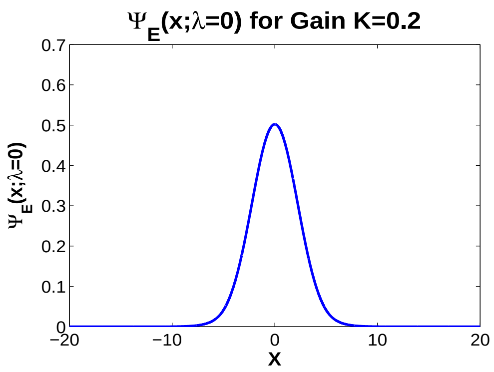
where u, $u' > 0$. We have completed the proof for Theorem 2.

Remark 6 Notice that we have used scaling here: $\frac{1}{\sqrt{\pi}}\Psi_E(x; u)$ and $\frac{1}{\sqrt{\pi}}\Psi_O(x; u)$ are the eigenfunctions satisfying the normalization condition (2.36).

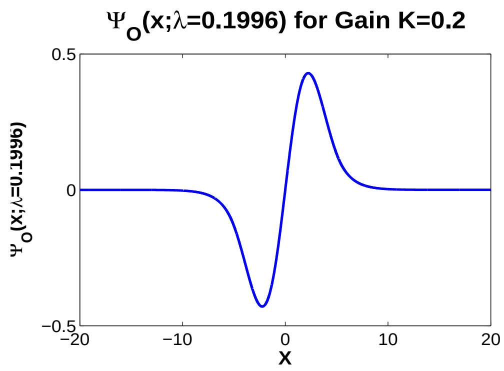
B4: Projection to Find the Conditional PDF From the projection coefficient formulas (2.38) and (2.39), it is easy to determine that the projections of the $e^{U(x')}\delta(x - x')$ on the discrete and continuous eigenfunctions are

$$\begin{aligned}
B(i) &= e^{U(x')}\Psi_D(x'; i) \\
B_E(u) &= e^{U(x')}\Psi_E(x'; u) \\
B_O(u) &= e^{U(x')}\Psi_O(x'; u).
\end{aligned}$$

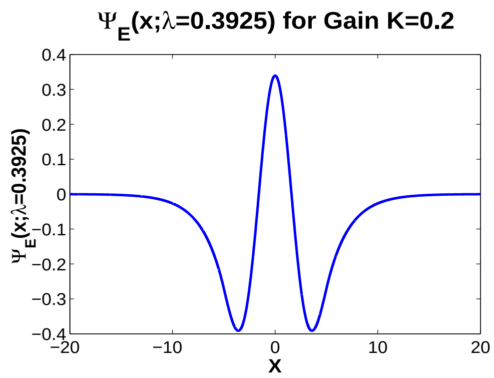
From the synthesis equation (2.32), we immediately get the conditional PDF (2.26). Thus, we have completed the proof for Theorem 1.



(a) Eigenfunction for $\lambda_0 = 0$



(b) Eigenfunction for $\lambda_1 = 0.1996$



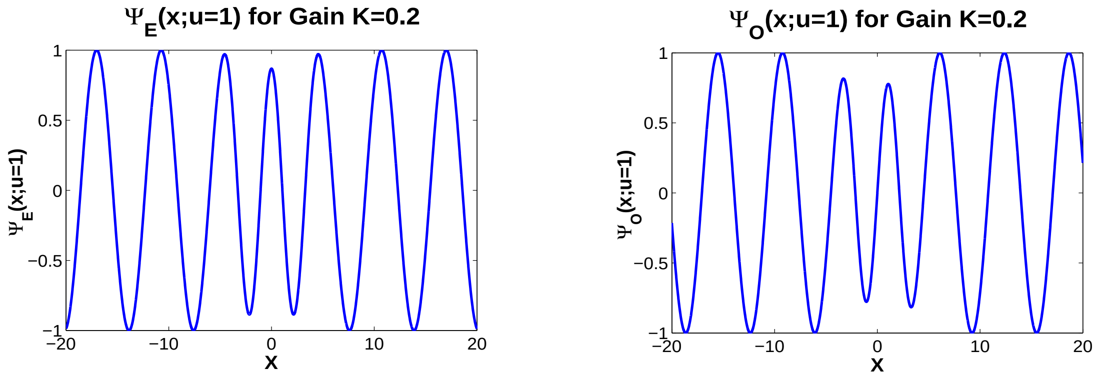
(c) Eigenfunction for $\lambda_2 = 0.3925$

Fig. 2.4: Eigenfunctions of the Schrödinger Operator associated with the Discrete Eigenvalues

2.3.4 Examples

In this section, we illustrate our analysis with two examples. Specifically, choosing the **feedback gain** $K = 0.2$ in (2.25), we compute $P(x, t; x' = 0, t_0 = 0)$ and $P(x, t; x' = 2, t_0 = 0)$.

Let us begin by finding the eigenfunctions. For $K = 0.2$, we find there are only 3 eigenvalues $\lambda < \frac{1}{2}$ that have decreasing exponential eigenfunctions, namely, $\lambda_0 = 0$, $\lambda_1 = 0.1996$ and $\lambda_2 = 0.3925$, whose normalized eigenfunctions are shown in Figure 2.3.4.



(a) Even Eigenfunction for $\lambda = 1$

(b) Odd Eigenfunction for $\lambda = 1$

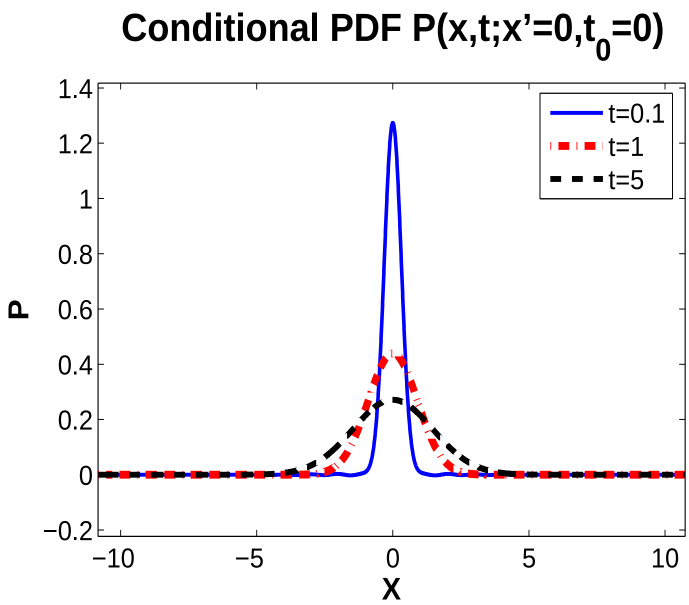
Fig. 2.5: Even and Odd Eigenfunctions of the Schrödinger Operator associated with $\lambda = 1$

For $\lambda > \frac{1}{2}$, each eigenvalue has both an even eigenfunction and an odd eigenfunction. One example of the normalized eigenfunctions in this case is illustrated in Figure 2.3.4.

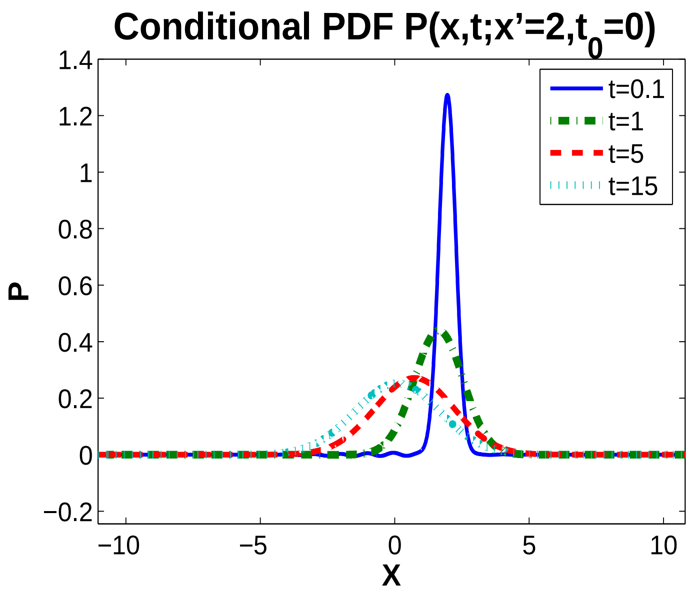
From the eigenfunctions, we can compute the conditional PDFs $P(x, t; x' = 0, t_0 = 0)$ and $P(x, t; x' = 2, t_0 = 0)$ according to (2.26) (where the integral is done numerically). The conditional PDFs at some specific times are illustrated in Figure 2.6.

Figure 2.6 demonstrates both drift and diffusion of the state from a specified initial condition ¶. The simulation result illustrates clearly the process of "diffusion", i.e. the increment of $V[x(t)]$ with time and if $x' \neq 0$, the process of "drift", i.e. $E[x(t)]$ approaches 0 with time. We also notice that in these two cases, we can regard $x(t)$ is being in steady-state for $t \geq 15$.

¶The slight glitches in Figure 2.6 are due to numerical implementations.



(a) Conditional PDF $P(x, t; x' = 0, t_0 = 0)$



(b) Conditional PDF $P(x, t; x' = 2, t_0 = 0)$

Fig. 2.6: Conditional PDF $P(x, t; x' = 0, t_0 = 0)$ and $P(x, t; x' = 2, t_0 = 0)$

2.4 Conclusion and Future Work

In this Chapter, we have advanced the understanding of the external responses of Saturating Static-Feedback-Controlled Systems by studying two canonical problems. In addition to the importance of the two problems per se, studying these two problems have also several important implications. The study of the deterministic disturbance of a Saturating Double-Integrator indicates the magnitude and structure of the external disturbances contribute to the complexity of the external stability concept of nonlinear systems. The study of the stochastic response of a Saturating Single Integrator driven by white noises implies the complexity of achieving the dynamic PDFs of nonlinear systems driven by stochastic processes, and in some sense, indicates some other methods, such as characterizing the variance of the states, should be applied for more complicated nonlinear systems.

There exist many directions for future works in this field and we list two most direct ones: the first is to study the deterministic external disturbance responses for general saturating feedback-controlled systems; the second is to study the dynamic response of a saturating feedback-controlled double integrator driven by white noise, and develop the optimal control law for it.

Chapter 3

On the Decentralized Control of Communicating-Agent Networks with Security Constraints

3.1 Introduction

Distributed algorithms and controllers that exploit a network's topological structure to complete cooperative tasks are of wide interest in several domains. For instance, autonomous-vehicle teams require controllers for formation flight, as well as algorithms for e.g. agreement and distributed partitioning tasks [21, 35]. Analogously, sensor networks may require a family of distributed algorithms e.g. for data fusion and fault recovery. We put forth the perspective here that *security* concerns—the need for protecting critical aspects of network computations/dynamics from adversaries—are a fundamental problem that is common to network algorithm/controller design. We thus introduce a control-theoretic framework for studying security of both network algorithms and controllers.

Let us give context for the security problem studied here from two perspectives, namely a computer scientist's perspective and a control theorist's perspective. In the computer science community, there is an extensive literature on protecting network *components* from certain types of attacks, e.g. denial-of-service attacks [22]. *Encryption*, or protection of communications through coding, is also widely studied [23]. Here, we pursue the complementary task of designing the algorithm/controller dynamics and communication topology to provably maintain security *at a network-wide level* (while still achieving the desired computation/control task). Literature on this algorithm-design approach to security is much more sparse, but some important results have been obtained under the heading of *secure multiparty computation* (or simply multiparty computation) [37]. Specifically, the efforts on multiparty computation show how a group of (possibly semi-honest or dishonest) agents can compute a function of their initial opinions, in such a manner that they do not gain any further information about the other agents' opinions beyond what can be obtained from the function value. We also consider securing distributed algorithmic computations, in the case where an adversary observes a subset of the agents over time (which is akin to a semi-honest agent case in the multiparty computation literature). However, we approach the problem from a complementary control-theoretic viewpoint, with the motivation that security is deeply connected with the control-theoretic notions of observability and detectability. We believe that this control-theoretic approach is promising for securing modern network algorithms/controllers, because it is matched with the paradigm for such networks: that of simple but highly constrained agents with general interaction/communication topology cooperating on a complex dynamic task. We will show that the control-theoretic approach allows us to secure a class of linear algorithms/controllers (which are suited for networks with simple agents) for function-computation and dynamic tracking tasks, under quite general conditions on the communication topology for the agents. Using the control-theoretic framework, we will also formalize the notion of securing certain statistics of the network state rather than the entire state.

From the perspective of a control theorist, the exact/asymptotical determination of the information (i.e. **state** in a control theorist's terminology) of the control systems is of great importance for the controller design and there are many classical notions (observability, detectability, and filtering) and methodologies (observer and filter) associated with it. Recently, the requirements to achieve the design goals and secure the important information/statistics simultaneously in many applications have attracted the attention of the control theorists. These requirements to conceal rather than explore some information are additional constraints on the control-law design and can be formulated as *security* tasks (requirements, constraints). Especially, the increasing study on the decentralized-control systems in the control community motivates the necessity to propose the notions of security in the context of the decentralized-control systems. This is not only because the security requirements are part of the inherent design specifics for the decentralized-control systems in many cases, but also due to the fact that the existence of the security constraints sometimes significantly impact the design of controllers/algorithms. We also notice that one advantage of the decentralized controllers/distributed algorithms over the centralized controller/algorithm is that it may help to achieve the security requirements. Although a wealth of work have been done in the field of decentralized-control systems, however, to the best of our knowledge, there still do not exist proper notions of security for the decentralized-control systems. In this part, we motivate the notions of security in the framework of a classical decentralized-control system, a Communicating-Agent Network (CAN)– a network in which the agents collaborate to complete some specified tasks, as a first step toward exploring the concept of security for the general decentralized-control systems. Specifically, we first formulate the different notions of security in the context of CANs and present some preliminary results for two classical CANs– Single-Integrator Network (SIN) and

Double-Integrator Network (DIN).

3.2 Problem Formulation

In this section, we motivate and describe the **notions of security** for **Communicating Agent Networks (CANs)**. Specifically, we first conceptualize and classify the notions of security in the context of a general CAN in Subsection 3.2.1. We then introduce two canonical CANs– the Single Integrator Network (SIN) and Double Integrator Network (DIN) and separately review two classical tasks of common interest for CANs, namely tracking and agreement, in Subsection 3.2.2. The specific analysis/design problems to be solved in this paper– namely, analysis and design of secure SINs/DINs for tracking and agreement– are developed in Subsection 3.2.3. Section 3.2.4 contains two illustrative examples.

3.2.1 Notions of Security Associated with CANs

In layman's terms, the word *security* refers to the protection of a system from "danger" or "loss". Variants of this notion have been introduced in different fields, such as coding theory, computer networking and communications. For a CAN pursuing a collaborative task, we always associate the notions of "source" and "goal" with the notion of security. The source is the information an adversary can eventually observe/violate, and the goal is the information/statistics that must be protected from the adversary. The CAN is said to be secure if the adversary cannot figure out the goal from the source. We will classify particular definitions of security according to the specific choices of source and goal.

Before conceptualizing specific notions of security in the context of a CAN, we first briefly present the model of a general CAN: For an arbitrary CAN with n **agents**, labeled $1, \cdots, n$, we assume that each agent i has a vector **state** $\mathbf{x_i(t)}$ that can be driven by a (vector) **actuation** signal $\mathbf{u_i(t)}$. We also assume each agent has some sensing capabilities, and hence has a (vector) **observation** signal $\mathbf{y_i(t)}$ that in general *depends on the current/past states of some or all of the agents.* In order to cooperatively complete certain dynamic task, each agent i in a CAN uses a *controller/protocol* that generates its actuation signal from its observation, so as to achieve desirable internal dynamics (move its state to a desirable value). We stress here that the agents' dynamics are uncoupled until the observations are used by the feedback controller/protocol, and hence design of the controller/protocol is critical for achieving security. For a given **agent set** S, we use the notations $\mathbf{x_S(t)}$, $\mathbf{u_S(t)}$ and $\mathbf{y_S(t)}$ to denote the vectors compromising the states, actuations and observations of all the agents in S *. For convenience, we use $\mathbf{x(t)}$, $\mathbf{u(t)}$ and $\mathbf{y(t)}$ to denote collections of the states, actuations, and observations of all agents, respectively.

Statistics in this form include the states of individual agents, state differences between agents, average states of multiple agents, etc. Very often, it is reasonable to assume that some set of agents S has been violated by an adversary (i.e. the "source" is an agent set S), and that either the actuations or observations or both of the agents in S are disclosed. Further, we contend many important statistics that must be protected from the adversary (the "goal") can be expressed as a linear function of the overall state $x(t)$. We also make the worst-case assumption that the adversary has knowledge of the **system model**, i.e. the model for the internal dynamics, observation topology, and controller/protocol of each agent. We thus consider the following definitions of **(S,M) Security** and **Overall Security** of a CAN, which are defined as follows:

Definition 2 ***(S,M) Security and Overall Security*** *Consider a CAN for which the observation*

*WLOG, we place the states/actuations/observations in increasing order with respect to the agent labels.

(or actuation) of an agent set S has been violated, and so an adversary has observed $\mathbf{y_S}$ ($\mathbf{u_S}$) over an interval $[t_1, t_2]$,

- *If the adversary cannot uniquely compute any statistic of the form $\mathbf{z(t)} = \mathbf{m^T x(t)}$, for any **nonzero** $\mathbf{m}$ in the range space of a matrix M and for every time $t \in [t_1, \infty)$, then we say that the CAN achieves **observation (actuation) (S,M) security**.*

- *If for every time $t \in [t_1, \infty)$ the adversary cannot uniquely determine $\mathbf{x(t)}$, then we say that the CAN achieves **overall observation (actuation) security** with respect to set S.*

(S,M) Security and Overall Security are also referred as **Exact (S,M) Security** and **Exact Overall Security**. As we shall illustrate through examples, (S,M) Security and Overall Security can sensibly capture typical security problems associated with a CAN. At the same time (as we shall show in the next section), both (S,M) Security and Overall Security can be related to the classical notion of *observability* in linear system theory, and hence some handy criteria can be developed to test them. We note that the overall security is *weaker* than any particular (S,M) Security requirement. It is also worth noting that some security conditions of interest are of the following form: if anyone of several (S,M) security requirements are met, then the CAN is said to be secure; the analysis have readily generalizes to this case.

The notions of (S,M) Security and Overall Security are still too *weak* for some applications, in the sense that the adversary may be able to approximate important information that should be protected (goals) from the disclosed information (sources), even though the exact computation of the state is impossible and hence (S,M) Security and Overall Security are achieved. Of particular interest, for a CAN in which the state of each agent i is required to follow a persistent *planned/desired trajectory* $\mathbf{\bar{x}_i(t)}$, it is possible that this planned/desired trajectory can be estimated even the state cannot. Notice that if the controller for each agent i is properly designed, then the real trajectory $\mathbf{x_i(t)}$ will approach the planned trajectory $\mathbf{\bar{x}_i(t)}$ after sufficient time has passed, and so the knowledge of the planned trajectory effectively give knowledge of the actual state trajectory in *steady-state*, or in other words, after sufficient time has passed. This possibility makes clear the stronger requirement of protecting the planned or steady-state trajectory is often important and motivates the following definitions:

Definition 3 ***Steady-State Security*** *Consider a CAN whose state* $\mathbf{x(t)}$ *approaches to a steady-state* $\mathbf{\bar{x}(t)}$*, and say that the observation (actuation) of an agent set* S *has been violated, and so an adversary has observed* $\mathbf{y_S}$ *(*$\mathbf{u_S}$*) over an interval* $[t_1, t_2]$*.*

- *If the adversary cannot uniquely compute any statistic of the form* $\mathbf{\bar{z}(t)} = \mathbf{m^T\bar{x}(t)}$*, for any* ***nonzero*** $\mathbf{m}$ *in the range space of a matrix* M *and* $t \in [t_1, \infty)$*, then we say that the CAN achieves the* ***steady-state observation (actuation) (S,M) security****.*

- *If the adversary cannot uniquely determine* $\mathbf{\bar{x}(t)}$ *at any time* $t \in [t_1, \infty)$*, then we say that the CAN achieves* ***overall steady-state observation (actuation) security*** *with respect to set* S*.*

This formulation of steady-state security clarifies the scenarios in which steady-state and basic notions of security should be used. In particular, we notice steady-state security implies that the adversary cannot determine a persistent aspect of the dynamics, while basic (S,M)/overall security only implies the transient. Specifically, the initial-condition information cannot be determined.

As we will see in the sequel, the above defined notions of security are all of interest in certain applications. To introduce the applications, we next concentrate on and apply the above-defined security notions to some canonical CANs.

3.2.2 Canonical CANs and Dynamic Tasks

In the remainder of the paper, we will apply the notions of security introduced above to some canonical CANs that are engaged in specific dynamic tasks of common interest. A range of networks of practical interest consist of agents with single- and double-integrator internal dynamics which make arbitrary linear observations. The single-integrator networks (SINs) and double-integrator networks (DINs) are expected to complete certain dynamic tasks cooperatively, two common tasks are tracking and agreement tasks (e.g. [29, 35]). Due to their extensive applicability, understanding and designing the security of these canonical networks is of particular importance. In this subsection, we carefully formulate the models of SIN and DIN, and the tracking and agreement task.

Single- and Double-Integrator Networks

Let us specify the SIN and DIN models by describing their internal dynamics and observation/sensing topology respectively.

Internal Dynamics: The internal dynamics of each agent in a Single Integrator Network (SIN) is described by $\dot{x}_i = u_i$, where x_i is the **state** of agent i and u_i is its **input/ actuation**. For convenience, we define a full state vector $\mathbf{x^T} = [x_1 \cdots x_n]$ and an input vector $\mathbf{u^T} = [u_1 \cdots, u_n]$. Similarly, the internal dynamics of each agent in a Double Integrator Network (DIN) is described by $\ddot{r}_i = u_i$, where r_i is the **position** of agent i and u_i is its **input/ actuation**. We note that r_i together with the **velocity** $s_i = \dot{r}_i$ specify Agent i's state. An full state vector is defined as $\mathbf{x^T} = [r_1 \cdots r_n,\ s_1 \cdots s_n]$ and an input vector same as that for a SIN.

Observation Topology: As mentioned above, each agent in a SIN/DIN has some sensing capabilities. Our model for sensing in SINs/DINs is quite general: we assume that each agent in a SIN/DIN makes multiple observations, each of which is a linear combination of the agents' current states. That is, for the SIN, each Agent i makes observations $y_i = G_i x$, where y_i is an m_i-component vector and we call the $m_i \times n$ matrix G_i the **topology matrix** for agent i. We also find it convenient to assemble the observations made by the agents into a single full observation vector $\mathbf{y^T} = \left[y_1^T \cdots y_n^T\right]$.

For the DIN, we assume that each agent makes identically-structured observations on positions and velocities. That is, Agent i has available two sets of observations

$$\begin{aligned} y_{ri} &= G_i r_i \\ y_{si} &= G_i s_i, \end{aligned}$$

where G_i is the **topology matrix** of Agent i. Similarly, a single full observation vector is defined as $\mathbf{y^T} = \left[y_{r1}^T \cdots y_{rn}^T,\, y_{s1}^T \cdots y_{sn}^T\right]$.

The observation topology can be effectively represented through a directed graph named **observation graph**, which is formulated as follows:

Definition 4 *The observation graph for a SIN/DIN is a* ***directed graph*** *such that*

1) *Each agent i in the SIN/DIN corresponds to a Vertex i in the observation graph;*

2) *There is a directed edge from Vertex i to Vertex k in the observation graph if and only if Agent k observes the state of Agent i, that is, the i-th column of G_k, which is henceforth denoted as $G_k(i)$, is a* ***nonzero*** *vector.*

Agent i is an **upstream neighbor** of Agent k if there exists a directed edge from Vertex i to Vertex k in the observation graph. Agent i is an **upstream neighbor** of Agent Set S if there exists a $k \in S$ such that Agent i is the upstream neighbor of Agent k. If Agent i is an upstream neighbor of Agent k, Agent k is called a **downstream neighbor** of Agent i. The **upstream agent set**

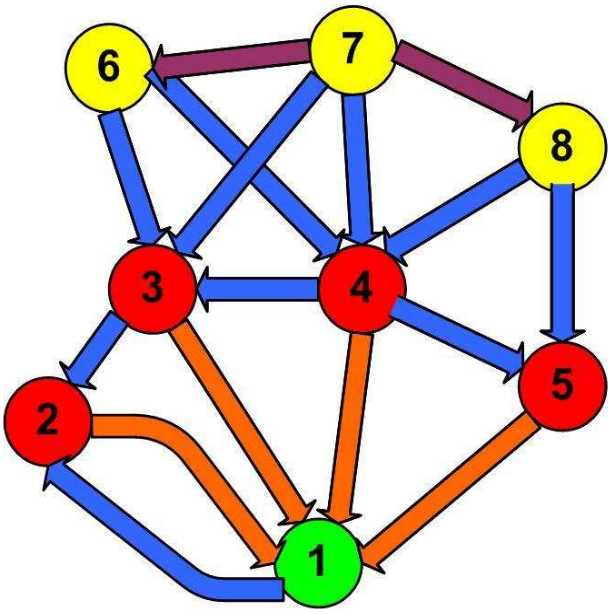


Fig. 3.1: Illustration of the Upstream Neighbors

$U(k)$ of Agent k is defined as the set of all the upstream neighbors of Agent k. The **upstream agent set** $U(S)$ of an agent set S is defined as $U(S) = \{i | \exists k \in S\, s.t.\, i \in U(k)\}$. Figure 3.1 presents an example of the observation graph and illustrates the notion of upstream neighbors.

Dynamic Task of SIN/DIN

We note the SIN/DIN models can represent various modern network dynamics, including such as the autonomous vehicle team dynamics or interactions among sensing agents for the purpose of data fusion (e.g. [26, 27, 35]). For these applications, SIN/DIN may need to collaborate to complete various dynamic tasks, such as tracking, formation, alignment, agreement, or distributed partitioning tasks , by using appropriate local controllers/protocols. The dynamic task that is required of the SIN/DIN motivates the type of security task (basic vs. steady-state) that must be considered. It also specifies the form of local controller, and hence impacts whether or not security can be achieved. In this subsection, we carefully formulate the tracking and agreement tasks, respectively.

Tracking Task for SIN/DIN: In many applications, the agents in a network must complete *tracking task*–ones in which each agent's state (i.e. position for double integrator) must follow a **reference signal (planned trajectory)** $\bar{x}_i(t)$. For example, teams of autonomous vehicles which sense each other's relative positions may need to follow a "lawnmower" pattern to search a target such as a landed satellite. Another example is that a group of missiles aimed at a target must follow some specific trajectories, to optimize the effectiveness of the attack. We note that more specific tasks, such as *formation* and *alignment* tasks, can be regarded as special cases of tracking task.

We can straightforwardly design controllers to make the agents in SIN/DINs track arbitrary reference signals, see [29]. However, in practice, nearly every trajectory of interest including sinusoidal, exponential and polynomial signals can be modeled as the output of an autonomous linear system with a proper initial condition. It is convenient for us to view the signal to be tracked as being generated from such an autonomous system, which is classically termed the *exosystem*. In particular, we assume that the command signal of Agent i can be generated as

$$\begin{cases} \dot{\omega}_i(t) = Q_i \omega_i(t) \\ \bar{x}_i(t) = d_i^T \omega_i(t) \end{cases},$$

and $\omega_i(t_0) \equiv \omega_{i,0}$.

For a SIN/DIN pursuing a tracking task, the security of the reference signal is often critical since it describes the persistent trajectories of agents and must be protected to safeguard the agents themselves. It is straightforward to verify that the security of the reference signal is in fact the steady-state security introduced above, hence we claim that the the steady-state security of SIN/DIN achieving tracking task is a canonical problem.

Agreement Task for SIN: In many applications involving multi-agent systems, such as a computer network, a sensor network and unmanned aerial vehicles, all the agents need to agree upon certain quantities of interest through a distributed averaging process, which is classically termed as *agreement tasks*. Specifically, each agent has an initial opinion (state) of the quantities and can update its own opinion using its current (single or multiple) observations, which are functions of the current opinions of all the agents. Agreement problem has a long history in automata theory and recently there are a lot of works solving it from control-theoretic perspective. One of the classical models for the agreement task from this viewpoint is SIN (see [35]). That is, each agent can update its opinion (state) using a first-order dynamics and a *static linear protocol* (i.e. a memoryless controller). The agreement task is achieved if there exists a static linear protocol such that the opinions (states) of all agents converge to the same (but in general initial opinion-dependent) value for all initial opinions (states). The eventual opinion reached by all the agents is defined as **agreement value**. Since both the protocol and the system dynamics are linear, a linear *agreement law* relating the initial condition and the agreement value can be developed.

For a SIN achieving the agreement task, the security of the initial opinions of some agent set is very important in many applications. For example, for a sensor network in which all the sensors measure the same target from different positions and an agreement task must be achieved to optimize the measurement, the initial opinion of one agent indicates the quality of the measurement made by that agent and requires to be protected from adversaries. Note that the security of the initial state can be achieved if some specific (S,M) security is achieved and hence can be viewed as a special case of (S,M) security.

In conclusion, we define the **structure** of a SIN/DIN as the collection of the observation topology and reference signal (for the tracking task) of all the agents. Note the structure of a

SIN/DIN can be modified in some applications but pre-set in others.

3.2.3 Security Task Associated with SIN/DIN

Motivated by the notions of security developed in Subsection 3.2.1 and the models and design tasks formulated in Subsection 3.2.2, we motivate and propose the specific analysis/design problems that will be solved in the remainder of the paper on the security tasks associated with SIN/DIN. The problems are formulated as follows:

Problem 1 ***Security and Observability*** *For a SIN/DIN with given structure and local controllers, what is the relationship between the notions of security and the observability/detectability of the closed-loop system?*

Problem 2 ***Security, Structure and Controller*** *For a SIN/DIN with given structure and local controllers, in what way do the structure and local controllers affect the achievability of the security? On the other hand, given the freedom of design, how can we design the local controllers and even structure to achieve the security?*

Before presenting some preliminary results on the above problems, we first demonstrate two examples in Subsection 3.2.4.

3.2.4 Examples

Some examples of the security problems associated with SIN/DINs achieving tracking/agreement tasks are presented below:

Example 1: A Team of Robots Searching for an Oil Field Assume an oil company is searching for an oil field in a desert. To avoid the exposure of human-beings to the extreme climate in the desert, a group of robots, labeled as Robot $1, 2, \cdots, n$, are used to finish this task. In order to efficiently complete the searching task, each robot, which can be modeled as a double integrator must follow a specific command signal. WLOG, each command signal can be modeled as a sinusoid with offset, that is

$$\begin{cases} x_i(t) = A_i sin(\omega_i t + \phi_i) + B_i \\ y_i(t) = t \end{cases},$$

where $x_i(t)$ and $y_i(t)$ are horizontal and vertical positions of Robot i. We also assume Robot 1 can measure its absolute position/velocity while other robots can only measure their relative positions/velocities with respect to its upstream neighbors. By adjusting its acceleration according to the received command signal and local measurement, after a short period of time, each robot approaches arbitrarily close to the desired trajectory specified by the command signal. Assume some spies want to prevent the completion of the searching task by attacking some robots, and unfortunately, they can violate some other robots by detecting their observations or actuations. Thus, in order to ensure the completion of the searching task, we must guarantee the command signals (steady-state trajectories) of some robots are secure given the violation of the observations or actuations of some other robots. That is, this problem is equivalent to a Steady-State Security of DIN achievng tracking tasks.

Example 2: The Agreement of Sensors Assume there are some limited number of sensors which collaborate to measure some target, such as the height of a mountain from different directions. Initially, each sensor has its own measurement available and uses it as its initial opinion. We further assume that each sensor also has an observation, which is a weighted average of the current opinions of its upstream neighbors and can update its opinion using first-order linear dynamics. Due to the limitation of physical condition of each sensor and the observation noise, the measurement of each sensor is not very accurate, and an optimal measurement can be achieved from a weighted average of the measurements of each sensor, which can only be achieved using agreement dynamics (algorithms) in each sensor. Assume some adversary want to figure out the conditions of some sensors by detecting those vehicles' own measurements (initial opinions), and unfortunately the observations of some other vehicles have been violated by the adversary. The security of the initial opinions of a "goal" set of sensors given the violation of a "source" set of sensors can be formulated as some special (S,M) security problem.

Although the above-proposed problems deal with very limited cases of the security problems associated with CAN, however, completely solving these problems are definitely a non-trivial task. In Section 3.3, several preliminary results for these problems are presented. Some future works are introduced in Section 3.4.

3.3 Preliminary Results

In this section, we present some preliminary results for the Security Tasks introduced in Section 3.2. Specifically, we first develop the **closed-loop models** for SIN/DIN achieving tracking tasks and SIN achieving agreement task in Subsection 3.3.1. Then, in Subsection 3.3.2, we explicitly present some test criteria for security and steady-state security (both (S,M) and overall) of SIN/DIN from the perspective of observability and detectability. At last, we explicitly study how the observation topology, structure of exosystem, and choice of local controller will affect the achievability of both the steady-state security and security, in Subsection 3.3.3.

3.3.1 The Closed-Loop Model of SIN/DIN

In this subsection, we will develop the closed-loop model for SINs/DINs achieving tracking/agreement tasks. Specifically, we develop the closed-loop model of SINs/DINs achieving tracking/agreement tasks separately and discuss their respective structures. For the convenience of formulating the conclusions, we define the **full topology matrix** as $G \equiv \begin{bmatrix} G_1 \\ \vdots \\ G_n \end{bmatrix}$.

For a SIN/DIN achieving the tracking task, first of all, we further delineate the eigenstructure of the exosystem. We find it convenient to assemble the multiple exosystems into a single exosystem. In particular, let us define the **exosystem state vector** as $w = [\omega_1 \cdots \omega_n]^T$, the **reference signal vector** $\bar{x} = [\bar{x}_1, \cdots, \bar{x}_n]^T$ and the **initial state vector** $w_0 = [\omega_{1,0} \cdots \omega_{n,0}]^T$. The exosystem state vector and the reference signal vector satisfy the state-space model

$$\begin{cases} \dot{\omega} = Q\omega \\ \bar{x} = D\omega \end{cases}$$

and $\omega(t_0) = \omega_0$, where $Q \equiv \begin{bmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{bmatrix}$ and $D \equiv \begin{bmatrix} d_1^T & & \\ & \ddots & \\ & & d_n^T \end{bmatrix}$.

Since the exosystem is our construction to model the reference signal, without loss of generality, we can assume (d_i^T, Q_i) is completely observable and Q_i is in the Jordan canonical form. Since d_i^T is a row vector (i.e. each exosystem has only one output), we can assume the eigenvalues associated with different Jordan blocks in Q_i are distinct †. We define the **Mode Set** Λ_i of the Agent i as the set of all eigenvalues of matrix Q_i.

†That is, the geometric multiplicity of each eigenvalue is 1. Notice this is not true if d_i^T is NOT a row vector, i.e. if the exosystem (3.1) has multiple outputs.

Closed-Loop Model of SIN/DIN Achieving Tracking Task: From our previous work [29], for a SIN/DIN achieving tracking task, we have the following lemmas:

Lemma 4 *Consider a SIN with given observation topology, exosystem, and decentralized control law*

$$u_i = k_i^T G_i(x - \bar{x}) + \dot{\bar{x}}_i, \tag{3.1}$$

for each Agent i*. This SIN achieves tracking task– i.e.* $\lim_{t\to\infty}(x(t) - \bar{x}(t)) = 0$ *for all initial conditions* $x(0)$ *– if and only if the eigenvalues of* KG *are in the OLHP, where* G *is the full topology matrix and* $K \equiv \begin{bmatrix} k_1^T & & \\ & \ddots & \\ & & k_n^T \end{bmatrix}$ *is the control* ***gain matrix****.*

Lemma 5 *Consider a DIN with given observation topology, exosystem, and decentralized control law*

$$u_i = \begin{bmatrix} k_i^T & \alpha k_i^T \end{bmatrix} \left(\begin{bmatrix} G_i & 0 \\ 0 & G_i \end{bmatrix} x - \begin{bmatrix} G_i D \\ G_i D Q \end{bmatrix} \omega \right) + d_i^T S_i^2 \omega_i, \tag{3.2}$$

for each Agent i*. This DIN achieves tracking task– i.e.* $\lim_{t\to\infty}(x(t) - \bar{x}(t)) = 0$ *for all initial conditions* $x(0)$ *– if and only if the eigenvalues of* KG *are in the OLHP, where* G *is the full topology matrix and* $K \equiv \begin{bmatrix} k_1^T & & \\ & \ddots & \\ & & k_n^T \end{bmatrix}$ *is the control* ***gain matrix****.*

One remark about the closed-loop system is worth mentioning:

Remark 5 *It is obvious that the system matrix* A *of the* ***closed-loop system*** *of SIN/DINs achiev-*

ing tracking task is

$$A = \begin{bmatrix} KG & -KGD + DQ \\ 0 & Q \end{bmatrix}, \tag{3.3}$$

for SIN and

$$A = \begin{bmatrix} 0 & I & 0 \\ KG & \alpha KG & -KGD - \alpha KGDQ + DQ^2 \\ 0 & 0 & Q \end{bmatrix} \tag{3.4}$$

for DIN.

Closed-Loop Model of SIN Achieving Agreement Task: A SIN is said to be governed by (or to have) the static linear protocol (K, z), where K is the block-diagonal matrix $K \equiv \begin{bmatrix} k_1^T & & \\ & \ddots & \\ & & k_n^T \end{bmatrix}$ and $\mathbf{z}$ is an n-component vector, if the actuation of Agent i is given by $u_i = k_i^T y_i + z_i$. Since the SIN has linear dynamics, if the agreement value for this network exists, it is an affine function of the agents' initial states, $\alpha = \mathbf{p}'\mathbf{x}(\mathbf{0}) + q$, where the pair $(\mathbf{p}, q)$ is defined as the **agreement law** for the network.

From our previous work [30], for a SIN achieving agreement task, we have the following lemma:

Lemma 6 *A SIN with observation topology matrix G and protocol $(K, \mathbf{z})$ reaches agreement if and only if either of the following two conditions hold:*

1) *All the eigenvalues of KG lie in the OLHP and $-(KG)^{-1}\mathbf{z} = \alpha\mathbf{1}$ for some α. In this case, the agreement law for the network is $(\mathbf{0}, \alpha)$. This scenario is defined as **Type 1 Agreement**.*

2) *KG has one eigenvalue of 0 with corresponding right eigenvector $\mathbf{1}$, the remaining eigenvalues of KG are in the OLHP, and $\mathbf{z} = \mathbf{0}$. In this case, the agreement law for the network is $(\mathbf{w}, 0)$*

*where w' is the left eigenvector of KG corresponding to the 0 eigenvalue. This scenario is defined as **Type 2 Agreement**.*

Remark 6 *Given the full topology matrix G and the protocol (K, z), we know the closed-loop model is*

$$\dot{x} = KGx + z. \tag{3.5}$$

3.3.2 Security Notions of SINs/DINs and Observability/Detectability

In this section, we explicitly develop the test criteria for both the **(exact) security** and **steady-state security** for a SIN/DIN from the perspective of notions of **observability** and **detectability** in classical linear system theory. Before presenting the test criteria, for an agent set $S = \{i_1, \cdots, i_m\}$, we find it convenient to define the **observation topology matrix for** S as $G_S = \begin{bmatrix} G_{i_1} \\ \vdots \\ G_{i_m} \end{bmatrix}$, **gian matrix for** S as $K_S = \begin{bmatrix} k_{i_1}^T & & \\ & \ddots & \\ & & k_{i_m}^T \end{bmatrix}$ and $D_S = \begin{bmatrix} d_{i_1}^T & & \\ & \ddots & \\ & & d_{i_m}^T \end{bmatrix}$. We note that $\mathbf{y_S} = G_S\mathbf{x}$. The test criteria for security and steady-state security of SIN are formulated in Theorem 7 and 8.

Theorem 7 *Consider a SIN with closed-loop system matrix A defined in (3.3), we have*

- *The SIN achieves the desired tracking task as well as the **observation (S,M) security** (**steady-state observation (S,M) security**) for an agent set S and a matrix M if and only if 1) the eigenvalues of KG are in the OLHP, and 2) $\left[M^T\, 0\right] V_{uo1}$ ($\left[M^T\, 0\right] V_{ud1}$) has full row rank, where V_{uo1} (V_{ud1}) is any basis of the **unobservable (undetectable) subspace** corresponding to the pair $([G_S\, 0]\,, A)$.*

- *The SIN achieves the desired tracking task as well as the* ***actuation (S,M) security*** *(****steady-state actuation (S,M) security****) for an agent set S and a matrix M if and only if 1) the eigenvalues of* KG *are in the OLHP, and 2)* $\left[M^T\, 0\right] V_{uo2}$ *(*$\left[M^T\, 0\right] V_{ud2}$*) has full row rank, where* V_{uo2} *(*V_{ud2}*) is any basis of the* ***unobservable (undetectable) subspace*** *corresponding to the pair* $\left(\left[K_S G_S,\, -K_S G_S D + D_S Q\right],\, A\right)$.

Proof:

We first prove the necessary and sufficient condition for the observation (steady-state) security. It has been shown that the tracking task can be achieved if and only if there exists a gain matrix K such that the eigenvalues of KG are in the OLHP. Hence, it remains to show that the SIN achieves the observation (S,M) security (steady-state observation (S,M) security) if and only if $\left[M^T\, 0\right] V_{uo1}$ ($\left[M^T\, 0\right] V_{ud1}$) has full row rank. Without loss of generality, we assume that $M \in R^{n\times r}$, where $r \leq n$.

We first prove the exact observation (S,M) security. Let $V_1, V_2, \ldots, V_j$ be the eigenvectors corresponding to unobservable modes of the system, and $V_{j+1}, \ldots, V_{j+l}$ be the eigenvectors corresponding to the observable modes. We note the observation (S,M) security is achieved if and only if for any nonzero $q \in R^r$ and $m^T = q^T M^T$, there exists an unobservable eigenvector V_k for the observability pair $\left(\left[G_S\, 0\right],\, A\right)$ such that $\left(m^T, 0\right) V_k \neq 0$. To see that, notice for any initial condition $\begin{bmatrix} \mathbf{x}(0) \\ \mathbf{w}(0) \end{bmatrix} = \sum_{i=1}^{n} \alpha_i V_i$, we know the statistics $z(t)$ is

$$z(t) = m^T x(t) = \left(m^T, 0\right) \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{w}(t) \end{bmatrix} = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} \left(m^T, 0\right) V_i, \tag{3.6}$$

where λ_i is the eigenvalue associated with eigenvector V_i. Note since $V_1, V_2, \ldots, V_j$ be the eigenvectors corresponding to unobservable modes of the system, $\alpha_1, \ldots, \alpha_j$ is not uniquely determined. Hence $z(t)$ can not be computed if there exists an unobservable eigenvector V_k such that $\left(m^T,0\right)V_k \neq 0$. On the other hand, the unique determination of $z(t)$ implies $\left(m^T,0\right)V_k = 0$ for any $1 \leq k \leq j$.

Notice $V_{uo} = [V_1, V_2, \ldots, V_j]$, thus the observation (S,M) security is achieved if and only if $\left(m^T,0\right)V_{uo} \neq \mathbf{0^T}$ for any $m^T = q^T M^T$ and $q \neq 0$. A little bit algebra implies this is equivalent to $(M^T,0)V_{uo}$ has full row rank.

The steady-state observation (S,M) security needs to consider the undetectable subspace V_{ud} of the observation pair instead of the unobservable subspace V_{uo}. The actuation (S,M) needs to consider a different observation pair. The proofs are very similar and we omit them here. □

Note the **overall security** (exact or steady-state, observation or actuation) is achieved when at least one state variable of the network cannot be computed by the adversary, which can be tested by considering multiple security requirements. Formally, we have the following theorem:

Theorem 8 *The SIN achieves the desired tracking task as well as the* ***overall security*** *(exact or steady-state, observation or actuation) if and only if 1) the eigenvalues of KG are in the OLHP, and 2) there exists at least one I_i such that $\left[I_i^T\, 0\right]V \neq 0$, where I_i is the ith column of an $n \times n$ identity matrix and V is any basis of the* ***unobservable subspace*** *(for exact security) or* ***undetectable subspace*** *(for steady-state security) corresponding to the pair $([G_S\, 0]\,,\, A)$ (for observation security) or $([K_S G_S,\, -K_S G_S D + D_S Q]\,,\, A)$ (for actuation security), where A is the closed-loop system*

Proof:

The proof follows immediately from the definition of overall security and the proof for (S,M) security.

□

Remark 7 *Similar conclusions can be derived for the security of DIN (exact or steady-state, (S,M) or overall, observation or actuation) by considering the unobservable (undetectable) subspaces as-*

sociated with proper observation pairs. For example, the proper observation pairs are

$$\left(\begin{bmatrix} G_S & 0 & 0 \\ 0 & G_S & 0 \end{bmatrix}, A\right), \quad \left(\begin{bmatrix} G_S & 0 & 0 \end{bmatrix}, A\right) \; or \; \left(\begin{bmatrix} K_S G_S & \alpha K_S G_S & -K_S G_S D - \alpha K_S G_S D Q + D_S Q^2 \end{bmatrix}, A\right)$$

respectively when all the observations (both position observations and velocity observations), position observations or actuation observations of an agent set S have been violated by an adversary.

Remark 8 *The test criteria formulated in Theorem 7 and 8 can be easily generalized to a general problem*

Problem 3 *Consider a LTI system model*

$$\begin{cases} \dot{x} = Ax \\ y = Cx \\ z = Dx \end{cases}, \tag{3.7}$$

given observation of y over an interval $[t_1, t_2]$, is the statistic z uniquely determined (can be exactly computed) over an interval $[t_1, \infty)$?

3.3.3 Security Notions of SINs/DINs and the Structures/Parameters of the Network

In the previous subsection, we have developed the explicit relationship between notions of security of SINs/DINs and notions of observability/detectability in the classical linear system theory and formulated very handy test criteria based on it. However, a SIN/DIN is a CAN with very special structure, hence we would like to explore how the structures and parameters of the network– i.e. **the observation topology**, **the local controllers/protocols** and **the eigenstructure of the exosystems** (for tracking task)– influence the achievability of the security task. This perspective not only leads to a new approach to test the achievability of the security of SINs/DINs, but more importantly, it illustrates a clear roadmap to achieve the proper security task by designing the structures and parameters of the network. Before presenting our main results, we first show that we can explore the **steady-state security** and the **security of the transients** separately, and then presents how the structures and parameters of SINs/DINs influence the achievability of both the steady-state security and exact security. Lemma 4 indicates we can consider the steady-state security and the security of the transients separately for a SIN achieving tracking task:

Lemma 7 *Consider a SIN achieving tracking task with closed-loop system matrix A defined in (3.3), we have*

1) *Assume λ is an eigenvalue of matrix Q_i and $v(i,\lambda)$ is the associated eigenvector, i.e. $Q_i v(i,\lambda) = \lambda v(i,\lambda)$. Thus there exists an eigenvector $\nu(i,\lambda)$ of A associated with λ and $\nu(i,\lambda) = \begin{bmatrix} D\mu(i,\lambda) \\ \mu(i,\lambda) \end{bmatrix}$, where $\mu(i,\lambda) = \begin{bmatrix} 0 \\ \vdots \\ v(i,\lambda) \\ \vdots \\ 0 \end{bmatrix}$ and $D\mu(i,\lambda) = \begin{bmatrix} 0 \\ \vdots \\ d_i^T v(i,\lambda) \\ \vdots \\ 0 \end{bmatrix}$. We define $\nu(i,\lambda)$ as the* ***augmented eigenvector*** *of $v(i,\lambda)$.*

2) *Assume λ is an eigenvalue of matrix KG and $v_G(\lambda)$ is the associated eigenvector, i.e. $KGv_G(\lambda) = \lambda v_G(\lambda)$ and $v_G(\lambda) \neq 0$. Thus the eigenvector of A associated with this particular λ is $\nu_G(\lambda) = \begin{bmatrix} v_G(\lambda) \\ 0 \end{bmatrix}$. We define $\nu_G(\lambda)$ as the* ***augmented eigenvector*** *of $v_G(\lambda)$.*

Proof:

First, define $\mu(i,\lambda) = \begin{bmatrix} 0 \\ \vdots \\ v(i,\lambda) \\ \vdots \\ 0 \end{bmatrix}$, notice since $Q_i v(i,\lambda) = \lambda v(i,\lambda)$, we have $Q\mu(i,\lambda) = \lambda\mu(i,\lambda)$.

Let $\nu(i,\lambda) = \begin{bmatrix} D\mu(i,\lambda) \\ \mu(i,\lambda) \end{bmatrix}$, we have

$$\begin{aligned} A\nu(i,\lambda) &= \begin{bmatrix} KG & -KGD+DQ \\ 0 & Q \end{bmatrix} \begin{bmatrix} D\mu(i,\lambda) \\ \mu(i,\lambda) \end{bmatrix} \\ &= \begin{bmatrix} KGD\mu(i,\lambda) - KGD\mu(i,\lambda) + DQ\mu(i,\lambda) \\ Q\mu(i,\lambda) \end{bmatrix} \\ &= \begin{bmatrix} \lambda D\mu(i,\lambda) \\ \lambda\mu(i,\lambda) \end{bmatrix} = \lambda \begin{bmatrix} D\mu(i,\lambda) \\ \mu(i,\lambda) \end{bmatrix} = \lambda\nu(i,\lambda) \end{aligned} \tag{3.8}$$

Thus, $\nu(i,\lambda)$ is an eigenvector of A associated with eigenvalue λ. Further notice

$$D\mu(i,\lambda) = \begin{bmatrix} d_1^T & & \\ & \ddots & \\ & & d_n^T \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ v(i,\lambda) \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ d_i^T v(i,\lambda) \\ \vdots \\ 0 \end{bmatrix}. \tag{3.9}$$

Assume λ is an eigenvalue of matrix KG and $v_G(\lambda)$ is the associated eigenvector and $\nu_G(\lambda) = \begin{bmatrix} v_G(\lambda) \\ 0 \end{bmatrix}$, we have

$$\begin{aligned} A\nu_G(\lambda) &= \begin{bmatrix} KG & -KGD+DQ \\ 0 & Q \end{bmatrix} \begin{bmatrix} v_G(\lambda) \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} KGv_G(\lambda) \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \lambda v_G(\lambda) \\ \lambda 0 \end{bmatrix} = \lambda \begin{bmatrix} v_G(\lambda) \\ 0 \end{bmatrix} = \lambda \nu_G(\lambda). \end{aligned} \quad (3.10)$$

This concludes the proof of Lemma 7.

Similar lemmas can be developed for other cases, such as a DIN achieving tracking task and a SIN achieving agreement task and we omit the derivation here.

Steady-State Security of SINs/DINs and the Structure of the Network

In this part, we are mainly interested in how the structures and parameters influence the steady-state security of SINs/DINs. As we will see in the following theorems, the achievability of security is affected by the observation topology and the eigenstructure of the exosystem, which are well connected by the following notion:

Definition 5 *For any eigenvalue $\lambda \in \cup_{i\in U(k)}\Lambda_i$, define $N_k(\lambda)$ as the number of agents in $U(k)$ whose mode set contains λ. We label these $N_k(\lambda)$ agents as Agent $k_1, \cdots, k_{N_k(\lambda)}$. Define the λ-**associated topology matrix** of Agent k as $G_k^\lambda = \left[G_k(k_1), \cdots, G_k(k_{N_k(\lambda)})\right]$. The λ-**associated topology matrix** of Agent Set S is defined similarly.*

The main results are formulated in the following theorems:

Theorem 9 *Consider a SIN and assume the security of the observation of Agent Set S has been violated, then the SIN achieves* ***steady-state observation (S,M) security*** *for $M = I_i^T$ if and only if either of the followings is true:*

1) *Agent i is NOT an upstream neighbor of Agent Set S.*

2) *Agent i is an upstream neighbor of Agent Set S, and there exists a $\lambda \in \Lambda_i$ such that $rank(G_S^\lambda) < N_S(\lambda)$.*

Similarly, assume the security of the actuation of Agent Set S has been violated, thus the SIN achieves ***steady-state actuation (S,M) security*** *for $M = I_i^T$ if and only if either of the followings is true*

1) *Agent i does NOT belong to Agent Set S*

2) *Agent i belongs to Agent Set S but 0 is an eigenvalue of Q_i.*

Proof:

We first prove the steady-state observation (S,M) security criteria and then prove the steady-state actuation (S,M) security criteria.

Proof for Steady-State Observation (S,M) Security

(Sufficiency) Notice in this case, the corresponding observation pair is $((G_S, 0), A)$. Consider an arbitrary eigenvector $v(i,\lambda)$ of Q_i and its augmented eigenvector $\nu(i,\lambda)$. From Lemma 7, we have

$$(G_S, 0)\nu(i,\lambda) = d_i^T v(i,\lambda) G_S(i),$$

thus, if Agent i is NOT an upstream neighbor of Agent Set S, we have $G_S(i) = 0$, thus, we have $(G_S, 0)\nu(i,\lambda) = 0$, which implies the mode λ of the reference signal (e.g. steady-state trajectory if the tracking task can be achieved) of Agent i, is NOT observable from the observation of Agent Set S. Thus, the adversary cannot figure out the reference signal of Agent i, thus Agent i is steady-state secure. [‡]

On the other hand, assume Agent i is an upstream neighbor of Agent Set S and there exists a $\lambda \in \Lambda_i$ such that $rank(G_S^\lambda) < N_S(\lambda)$. This implies that there exist $N_S(\lambda)$ upstream neighbors of Agent Set S whose reference signals contain mode λ, equivalently there exist $N_S(\lambda)$ linearly independent eigenvectors $\nu(i_1, \lambda), \cdots, \nu(i_{N_k(\lambda)}, \lambda)$ of A associated with eigenvalue λ such that $(G_S, 0)\nu(i_j, \lambda) \neq 0$ for $j = 1, \cdots, N_k(\lambda)$. Agent i is secure if the eigenvalue λ is NOT completely observable for the observation pair $((G_S, 0), A)$, that is, if there exists an eigenvector of A associated with λ, which can be expressed as a non-trivial linear combination $\sum_{j=1}^{N_k(\lambda)} \alpha_j \nu(i_j, \lambda)$ such that $(G_S, 0) \sum_{j=1}^{N_k(\lambda)} \alpha_j \nu(i_j, \lambda) = 0$. Using the conventions of Lemma 2, we have

$$(G_S, 0) \sum_{j=1}^{N_k(\lambda)} \alpha_j \nu(i_j, \lambda) = \sum_{j=1}^{N_S(\lambda)} \alpha_j d_j^T v_j G_S(i_j) = 0. \tag{3.11}$$

Since $d_j^T v_j \neq 0$ for any $1 \leq j \leq N_S(\lambda)$ and not all α_j's are 0, we know that the vectors $G_S(i_j)$ for $j = 1, \cdots, N_S(\lambda)$ are linearly dependent. Thus we have $rank(G_S^\lambda) < N_S(\lambda)$.

(Necessity) Assume the reference signal of Agent i is secure given the violation of the observations of Agent set S. We consider 2 possible cases:

1) Agent i is NOT an upstream neighbor of Agent set S. From the proof of the Sufficiency, we know Agent i is secure.

2) Agent i is an upstream neighbor of Agent set S, we prove the Case 2 must be true by contradiction. If not, for any $\lambda \in \Lambda_i$, we have $rank(G_S^\lambda) = N_S(\lambda)$. So any eigenvalue λ associated with Q_i is completely observable for the observation pair $((G_S, 0), A)$. Thus, the

[‡]In fact, this is true for any $\lambda \in \Lambda_i$, that is every mode of the reference signal of Agent i can not be determined.

reference signal of Agent i is NOT secure.Contradiction.

Proof for Steady-State Observation (S,M) Security

Notice in this case, the observation pair is $\big((k_i^T G_i,\, -k_i^T G_i D + d_i^T Q),\, A\big)$. Assume v_j is an eigenvector of Q_j associated with eigenvalue λ, and ν is the corresponding augmented eigenvector of matrix A associated with λ. We have

$$\begin{aligned}(k_i^T G_i,\, -k_i^T G_i D + d_i^T Q)\nu &= k_i^T G_i d_j^T v_j - k_i^T G_i d_j^T v_j + d_i^T Q_i v_j \\ &= d_i^T Q_i v_j \end{aligned} \tag{3.12}$$

Note if $i \neq j$, we have $d_i^T Q_i v_j = 0$; if $i = j$, we have $d_i^T Q_i v_i = \lambda d_i^T v_i$. Since $d_i^T v_i \neq 0$, we have $d_i^T Q_i v_i \neq 0$ as long as $\lambda \neq 0$.□

Similarly, we can prove the corresponding conclusions for DINs in Theorem 10 and 11. We omit the proofs for these theorems.

Theorem 10 *Consider a DIN and assume the security of the position observation of Agent Set S has been violated, then the DIN achieves **steady-state observation (S,M) security** for $M = \left(I_i^T, 0\right)$ and $M = \left(0, I_i^T\right)$ if and only if either of the followings is true*

1) Agent i does NOT belong to Agent Set S

2) Agent i belongs to Agent Set S but 0 is an eigenvalue of Q_i.

Theorem 11 *Consider a DIN and assume the security of the velocity observation of Agent Set S has been violated, then the DIN achieves **steady-state observation (S,M) security** for $M = \left(I_i^T, 0\right)$ if and only if any of the followings is true*

1) Agent i is NOT an upstream neighbor of Agent Set S.

2) Agent i is an upstream neighbor of Agent Set S, and there exists a $\lambda \in \Lambda_i$ such that $rank(G_S^\lambda) < N_S(\lambda)$;

3) The reference signal of Agent i contains a DC mode.

And the DIN achieves ***steady-state observation (S,M) security*** *for $M = \left(0, I_i^T\right)$ if and only if either of the followings is true*

1) Agent i does NOT belong to Agent Set S

2) Agent i belongs to Agent Set S but 0 is an eigenvalue of Q_i.

Exact Security of SIN and the Decentralized Controller/Protocol Design

As we have disscussed before, the **exact security** (**(S,M)** or **overall**, **observation** or **actuation**) of a SIN/DIN can be achieved by either achieving the **steady-state security** or guaranteeing the **transients** of the **states** of the network are secure. In this part, we explore how the observation topology of a SIN affects the the transients of the states of the network.

We contend that for a SIN with arbitrary observation topology, to design a set of static decentralized controllers to achieve both the security of the transients and the dynamic task is nontrivial. However, for some SINs with classical observation topologies, many important conclusions can be developed. Specifically, in this part, we first show that the transients security problem of a SIN achieving either tracking or agreement task can be rephrased as a classical **KG Design** problem in decentralized control and then presents the results for some special observation matrices, such as the **D-stable matrix** and **Laplacian matrix**. First, we reformulate the problem as a KG Design problem:

Theorem 12 *A SIN can achieve the dynamic task (either tracking or agreement) and the overall* *observation (actuation) security if either of the followings is true*

1) *The SIN achieves the dynamic task and the overall steady-state observation (actuation) security.*

2) *For the observation topology matrix G of the SIN, if there exists a block diagonal matrix K such that (1) the eigenvalues of KG satisfy the conditions to achieve the dynamic task (in OLHP); (2) the observation pair (G_S, KG) $((K_S G_S, KG)$ for overall actuation security $)$ has an unobservable eigenvalue.*

Proof:

The dynamic task and the overall exact observation (actuation) security of a SIN can be achieved if and only if 1) the dynamic task and the overall steady-state observation (actuation) security are achieved, or 2) the dynamic task is achieved and the transients of the state are secure. It is obvious that the transients of the state are secure is equivalent to the condition that the observation pair (G_S, KG) (or $(K_S G_S, KG)$) has an unobservable eigenvalue. □

We are particularly interested in the overall observation (actuation) security of a SIN in which each agent has only one observation. Hence, the observation topology matrix G is square and the control gain matrix K is a diagonal matrix. Specifically, some important results can be developed for a SIN with special observation topology matrix G, such as the D-stable matrix and Laplacian matrix. The D-Stable Matrix is defined as follows:

Definition 6 ***D-Stable Matrix*** *A square matrix N is D-Stable if DN is stable for any positive diagonal matrix D.*

The Laplacian matrix is a widely used in graph theory and Graph-laplacian observation topologies are applicable when agents know their states (positions) relative to other agents.More specifically, each Agent i's observation is assumed to ba an average of differences between i's state and its upstream neighbors' states. Motivated by this idea, the Laplacian matrix is defined as follows:

Definition 7 *A SIN is said to have a Graph-Laplacian observation topology if its observation topology matrix G satisfies:*

$$G_{ij} = \begin{cases} |U(i)| & \text{if } i = j \\ -1 & \text{if } j \in U(i) \\ 0 & \text{otherwise} \end{cases},$$

where G_{ij} is the (i, j) component of G and $|U(i)|$ is the cardinality of $U(i)$.

Definition 8 *Two Scalars a and b are said to have the same sign pattern if either of the followings is true:*

1) $a = b = 0$

2) $ab > 0$

Two vectors v_1 and v_2 are said to have the same sign pattern if 1) v_1 and v_2 are of the same dimension; 2) the corresponding components of v_1 and v_2 have the same sign pattern.

Based on the above definitions, we have the following theorems:

Theorem 13 *Consider a SIN whose observation topology matrix G satisfies the condition that $-G$ is a D-stable matrix. Thus the SIN can achieve the dynamic task (either tracking or agreement) and the overall observation (actuation) security given the violation of Agent i by designing the static decentralized control law K if there exists a vector v such that*

1) The i-th component of v, v_i is 0 and $g_i^T v = 0$;

2) For $j \neq i$, $g_j^T v$ and v_j are 0 have the ***same sign pattern****.*

Proof:

Notice the dynamic task and the overall security can be simultaneously achieved if there exists a diagonal matrix K such that

1) $-K$ is a positive diagonal matrix, hence $KG = (-K)(-G)$ is stable;

2) $(g_i^T KG)$ $((k_i g_i^T KG)$ for actuation security) has an unobservable eigenvector, i.e. $KGv = \lambda v$ and $g_i^T v = 0$ ($k_i g_i^T v = 0$ for actuation security).

For the overall observation security, note since KG is stable and $-K$ is a positive diagonal matrix, we know all the eigenvalues of KG and all the diagoanl components of K are negative. Hence, from the equations

$$\begin{aligned} k_j g_j^T v &= \lambda v_j \quad \text{for any } 1 \leq j \leq n \\ g_i^T v &= 0 \quad \text{for Agent } i, \end{aligned} \tag{3.13}$$

we know $v_i = 0$ and for $j \neq i$, $g_j^T v$ and v_j have the same sign pattern.

For the overall actuation security, the proof is similar and we omit it here.□

Remark 9 *Similar conclusions can be derived for the case when G is the D-stable matrix by modifying the second condition as: for $j \neq i$, either both $g_j^T v$ and v_j are 0 or they have the* ***opposite sign pattern****.*

It is worth noting that for any Laplacian matrix L, we have $-L$ is a D-stable matrix. Since the Laplacian matrix is symmetric, the observation graph for a SIN with graph-Laplacian observation topology can be represented using an undirected graph. More specifically, for a broad class of SINs whose observation topology matrices are Laplacian matrices of graphs with single node-cut, we have the following theorem:

Theorem 14 *A SIN whose observation topology matrix G is a Laplacian Matrix of a graph with a single node-cut can achieve the dynamic task (either tracking or agreement) and the overall observation (actuation) security given the violation of the cut-node by designing the static decentralized control law K.*

Proof:

We prove the conclusion by showing that there exists a vector v satisfying the condition of Theorem 13. Note since G is the Laplacian matrix of a graph with a single node-cut, without loss of generality, we can assume $G = \begin{bmatrix} L_A & 0 & L_{Al} \\ 0 & L_B & L_{Bl} \\ L_{Al}^T & L_{Bl}^T & l \end{bmatrix}$. We select $v = \begin{bmatrix} a \\ b \\ 0 \end{bmatrix}$, where a and b are of proper dimensions and 0 is a scalar. Note

$$
\begin{aligned}
Gv &= \begin{bmatrix} L_A & 0 & L_{Al} \\ 0 & L_B & L_{Bl} \\ L_{Al}^T & L_{Bl}^T & l \end{bmatrix} \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} L_A a \\ L_B b \\ L_{Al}^T a + L_{Bl}^T b \end{bmatrix}. \qquad (3.14)
\end{aligned}
$$

Notice $L_{Al}^T a + L_{Bl}^T b$ is a scalar, hence we can always select an eigenvector v_A of L_A and an eigenvector v_B of L_B such that $L_{Al}^T v_A + L_{Bl}^T v_B = 0$ (Note this is generally not true if $L_{Al}^T a + L_{Bl}^T b$ is not a scalar.). From the properties of Laplacian matrix, we know $L_A v_A$ and v_A, $L_B v_B$ and v_B have the same sign pattern, respectively. Hence the vector $v^* = \begin{bmatrix} v_A \\ v_B \\ 0 \end{bmatrix}$ satisfies the condition of Theorem 13. □

3.4 Conclusion and Future Work

In this chapter, we have formulated the notions of security in the context of the Decentralized-Controlled Communicating-Agent Network and presented some preliminary results on SINs/DINs by relating the security notions to observability/detectability and structures of the networks. The future work has two directions: the first is to further the research on the relationship between the structures of SIN/DINs and the achievability of the exact security, and use some more advanced techniques such as the multiple-delay control laws to achieve it; the second is to further the problem by taking the disturbances and noises into consideration, and change the criteria from a boolean function to a more sensible measure of "degree of security".

Chapter 4

Conclusion

In this thesis, we have mainly focused on two specific topics in the field of linear systems with constraints, the External Stability of Linear Systems with Actuator Saturation Constraints and the Decentralized Control of Communicating-Agent Networks with Security Constraints. In the first topic, we further the research by studying two canonical problems– the Deterministic Disturbance Response of a Saturating Static-Feedback-Controlled Double Integrator and the Dynamic Response of a Saturating Static-Feedback-Controlled Single Integrator driven by white noise. In the second topic, we formulate the notions of security in the context of the Decentralized-Controlled Communicating-Agent Networks, and present some preliminary results on some canonical networks, i.e. SINs and DINs.

It should be noticed that these two problems are parts of a continuous efforts to better understand the impact of all kinds of constraints on linear systems. In fact, the Actuator Saturation and Decentralized-Control Constraint are two very common and important constraints in practice. It should also be pointed out that these two kinds of constraints can be easily combined, for example, each agent in a DIN, which is modeled as a double integrator, is typically subject to actuator saturation and external disturbances, which can be either deterministic or stochastic, depending on the specific problems. A decentralized-controlled Communicating-Agent Network with actuator saturation and external disturbances is a more sensible and practical model in many applications and research on it is a good direction in the future.

Bibliography

[1] A. Saberi, Z. Lin, and A. Teel, "Control of linear systems with saturating actuators", *IEEE Trans. Automatic Control*, vol. 41, no. 3, pp. 368-378, 1996.

[2] H. J. Sussmann, E. D. Sontag, and Y. Yang, "A general result on the stabilization of linear systems using bounded controls", *IEEE Trans. Automatic Control*, vol. 39, pp. 2411-2425, Dec 1994

[3] Z. Lin, A. Saberi, and A. R. Teel, "Simultaneous L_p-stabilization and internal stabilization of linear systems subject to input saturation– State feedback case", *System Control Letter*, vol. 25, no. 3, pp. 219-226, 1995.

[4] A. Saberi, P. Hou, and A. A. Stoorvogel, "On Simultaneous Global External and Global Internal Stabilization of Critically Unstable Linear Systems with Saturating Actuators", *IEEE Transactions on Automatic Control*, vol. 45, no. 6, June 2000. .

[5] G. Shi and A. Saberi, "On the input-to-state stability (ISS) of a double integrator with saturated linear control laws", *Proceedings of the American Control Conference*, May 2002.

[6] A. A. Stoorvogel, G. Shi, and A. Saberi, "External stability of a double integrator with saturated linear control laws". *Dynamics of Continuous, Discrete and Impulsive Systems*, Vol. 11, No.4-5, pp. 429-451, August 2004.

[7] Y. Chitour, "On the L Stabilization of the Double Integrator Subject to Input Saturation", *ESAIM: Control, Optimisation and Calculus of Variations*, March 2001, Vol. 6, 291-331.

[8] E. Sontag, "Input to State stability: Basic concepts and results", *Springer Lecture Notes in Mathematics*, 2005.

[9] H. Fang, Z. Lin and T. Hu, "Analysis of Linear Systems in the presence of Actuator Saturation and L_2 Disturbances", *Proceedings of the 42nd IEEE Conference on Decision and Control* , December 2003.

[10] M. Nikolaou and V. Manousiouthakis, "A Hybrid Approach to Nonlinear System Stability and Performance", *AIChE Journal*, April 1989, Vol. 35, No. 4 2005.

[11] Z. Wen, S. Roy and A. Saberi, "On the Disturbance Response and External Stability of a Saturating Static-Feedback-Controlled Double Integrator", accepted by *Automatica*.

[12] A. T. Fuller, "Analysis of nonlinear stochastic systems by means of the Fokker-Planck equation", *Int. J. Contr.,* vol. 9, pp. 603-655, 1969.

[13] H. Risken, "The Fokker-Planck Equation, methods of solution and applications", second edition, published by Springer-Verlag, 1989.

[14] R. W. Brockett and D. Liberzon, "On explicit steady-state solutions of Fokker-Planck equations for a class of nonlinear feedback systems", *Proceedings of the american Control Conference,* pp. 264 268, 1998.

[15] A. A. Stoorvogel, S. Weiland and A. Saberi, "On stabilization of linear systems with stochastic disturbances and input saturation", *43rd IEEE Conference on Decision and Control,* 2004.

[16] C. Cohen-Tannoudji, B. Diu, F. Laloe, *Quantum Mechanics, Volume 1*, Wiley, New York, 1977.

[17] C. Cohen-Tannoudji, B. Diu, F. Laloe, *Quantum Mechanics, Volume 2*, Wiley, New York, 1977.

[18] W. P. Yuen, " Exact analytic analysis of finite parabolic quantum wells with and without a static electric field", *Physical Review B*, Volume 48, Number 23, Dec. 15, 1993.

[19] W. J. Rugh, *Linear System Theory*, 2^{nd} edition, Chapter 3, Prentice Hall, 1996.

[20] S. H. Wang and E. J. Davison, "On the stabilization of the decentralized control systems", *IEEE Transactions on Automatic Control,* vol. AC-18, no. 5, pp. 473-478, 1978.

[21] Y. Wan, S. Roy, and B. C. Lesieutre, "A stochastic automaton-based algorithm for flexible and distributed network partitioning," submitted to the *IEEE Transactions on Parallel and Distributed Systems* .

[22] A. D. Wood, J. A. Stankovic, "Denial of service in sensor networks", *Computer*, Volume: 35, Issue: 10, pp. 54-62.

[23] R. M. Needham, M. D. Schroeder, "Using encryption for authentication in large networks of computers", *Communications of the ACM*, Volume 21 , Issue 12, pp. 993 - 999.

[24] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations", submitted to *IEEE Transactions on Automatic Control*, April 2003.

[25] S. Roy, A. Saberi, and K. Herlugson, 'Formation and Alignment of Distributed Sensing Agents with Double-Integrator Dynamics', IEEE Press Monograph on *Sensor Network Operations*, 2004 (in press).

[26] A. Williams, G. Lafferriere, and J. J. P. Veerman, "Stable motions of vehicle formations," in *Proceedings of the 2005 Conference on Decision and Control*.

[27] P. Ogren, E. Fiorelli, and N. Leonard, "Cooperative control of mobile sensor networks," *IEEE Transactions on Automatic Control*, vol. 49, no. 8, 2004.

[28] D. J. Stilwell, B. E. Bishop, and C. A. Sylvester, "Redundant manipulator techniques for partially decentralized path planning and control of a platoon of autonomous vehicles," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 35, no. 4, Aug. 2005.

[29] S. Roy, L. Chen, and A. Saberi, "On the Information Flow Required for Tracking Control in Networks of Mobile Sensing Agents", *ALGOSENSORS 2006:* 208-216.

[30] S. Roy, K. Herlugson, and A. Saberi, "A control-theoretic perspective on distributed discrete-valued decision-making in networks of sensing agents," to appear in the *IEEE Transactions on Mobile Computing*.

[31] A. Saberi, A. Stoorvogel, and P. Sannuti, *Output Regulation and Control Problems with Regulation Constraints*, Springer-Verlag, 1999

[32] R. O. Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions in Automatic Control*, vol. 49, pp. 1520-1533, Sep. 2004.

[33] M. E. Fisher and A. T. Fuller, "On the stabilization of matrices and the convergence of linear iterative processes," *Proceedings of the Cambridge Philosophical Society*, pp. 417–425, 1956.

[34] S. Roy, J. Minteer, and A. Saberi, "Some new results on stabilization by scaling," accepted to the *American Control Conference*, 2006.

[35] S. Roy, A. Saberi, and K. Herlugson, "A control-theoretic perspective on the design of distributed agreement protocols,", Submitted to the *Journal of the Association for Computing Machinery.*

[36] D. J. Stilwell, and B. E. Bishop, "Redundant manipulator techniques for path planning and control of a platoon of autonomous vehicles," *Proceedings of the 41st IEEE Conference on Decision and Control*, pp. 2093–2098, 2002.

[37] Y. Lindell, and B. Pinkas, "A prrof of Yao's Protocol for Secure Two-Party Computation", *electronic Colloquium on Computational Compexity*, Report No. 63, 2004.

[38] H. Kobayashi, and T. Y. Yoshikawa, "Graph-Theoretic Approach to Controllability and Localizability of Decentralized Control Systems," *IEEE Transactions on Automatic Control*, vol. 27, No. 5, Oct. 1982.

[39] D. Hershkowitz, "Recent direction in matrix stability," *Linear Algebra Appl.* 171 (1992), 161-186.