Chapter 1

INTRODUCTION

1.1 Background and motivation

ANNs, a new method for various information processing, is now widely used in the fields of pattern recognition, image processing, optimal computation, aerospace, defense, robotics, telecommunications, signal processing and combinatorial optimization, for details see [17, 30, 34]. NNs and their various generalizations have attracted the attention of the scientific community due to their ability to solve difficult optimization problems.

In hardware implementation of NNs, stochastic disturbances are nearly inevitable owing to thermal noise in electronic devices. It has been shown that certain stochastic inputs could destabilize NNs. On the other hand, time-delays are often encountered in various Engineering, biological and economic systems, which is frequently a source of oscillation and deterioration of the system. Recently, several sufficient conditions for the stability of SNNs with time-delays have been proposed. It should be noted that many unsolved problems still exist as far as the problems of existence, periodic and stability analysis for NNs with time-delays are concerned. Thus, the study of NNs with time-delays is an active research topic under considerable attention in recent years.
This chapter motivates the study of DNNs such as SNNs, neutral type NNs, Lyapunov stability and LMI. Further we shall give the occurrence of the various NNs in different application fields.

1.2 Introduction to neural networks

An ANN is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain and process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involve adjustments to the synaptic connections that exist between the neurons.

1.2.1 Artificial neural networks

The elementary nerve cell, called a neuron, is the fundamental building block of the biological NN. Its schematic diagram is shown in Figure 1.1. An artificial neuron is a computational model inspired in the natural neurons. Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron. When the signals received are strong enough (surpass a certain threshold), the neuron is activated and emits a signal through the axon. This signal might be sent to another synapse, and might activate other neurons.

The complexity of real neurons is highly abstracted when modelling artificial neurons. These basically consist of inputs (like synapses), which are multiplied by weights (strength of the respective signals), and then computed by a mathematical function which determines the activation of the neuron. Another function (which may be the
identity) computes the output of the artificial neuron (sometimes in dependance of a certain threshold). ANNs combine artificial neurons in order to process information. Its schematic diagram is shown in Figure 1.2.

For the general artificial neuron model, the contribution of each neuron to the system is determined by the weights. One of the first things of the model is to calculate the weighted sum of inputs. Weights can also be negative, so we can say that the signal is inhibited by the negative weight. Depending on the weights, the computation of the neuron will be different. By adjusting the weights of an artificial neuron we can obtain the output for the specific inputs. But when we have an ANN of hundreds or thousands
of neurons, it would be quite complicated to find by hand all the necessary weights. But we can find algorithms which can adjust the weights of the ANN in order to obtain the desired output from the network. This process of adjusting the weights is called learning or training. The number of types of ANNs and their uses are very high. The differences in them might be the functions, the accepted values, the topology, the learning algorithms, etc. Also there are many hybrid models where each neuron has more properties than the ones we are reviewing here. Because of matters of space, we will present only an ANN which learns using the backpropagation algorithm [78] for learning the appropriate weights, since it is one of the most common models used in ANNs, and many others are based on it. Since the function of ANNs is to process information, they are used mainly in fields related with it.

### 1.2.2 Neural network architectures

In any NN the neurons are grouped together into layers. Data is fed through the input layer, processed by one or several hidden layers and finally the network provides the output through the output layer.

- **Feed forward network:** In this kind of networks there is no feedback in the system. The input after processing through the hidden layers is directly given out as output.

- **Recurrent neural network architecture:** Feedback occurs in almost every part of the human brain. Thus a RNN was introduced and can be defined as a network that has feedback to all the inputs. The feedback can be given to all input neurons including the neuron’s own input.
1.3 Hopfield neural networks

One of the earliest RNNs reported in literature was the auto-associator independently described by Anderson [2] and Kohonen [45] in 1977. It consists of a pool of neurons with connections between each unit \( i \) and \( j \), \( i \neq j \). All connections are weighted. HNNs are recurrent networks introduced by Hopfield [37]. The earliest Hopfield network, which employs two-state (on/off) neurons, is used for the design of neural content addressable memories. Hopfield later introduced a modified version of his earlier model that employed a continuous nonlinear function to describe the output behavior of the neurons. The Hopfield network (model) consists of a set of neurons and a corresponding set of unit delays, forming a multiple-loop feedback system, as illustrated in Figure 1.3. The number of feedback loops is equal to the number of neurons. Basically, the output of each neuron is fed back via a unit delay element, to each of the other neurons in the network. In other words, there is no self-feedback in the network. The Hopfield network is useful as a content addressable memory or an analog computer for solving combinatorial-type optimization problems.

In 1982, Hopfield [37] brings together several earlier ideas concerning Hopfield networks and presents a complete mathematical analysis, see [1]. It is therefore that this network described in this chapter is generally referred to as the Hopfield network.

1.4 Implementation of Hopfield neural networks

A well-known model of dynamic RNNs with some useful collective computational properties is due to Hopfield. A continuous-time model of an analog NN can be described
Figure 1.3: Architectural graph of a Hopfield network consisting four neurons

by the following system of nonlinear DEs

State equation: \( C_i \frac{dx_i(t)}{dt} = -\frac{x_i(t)}{R_i} + \sum_{j=1}^{n} w_{ij} y_j(t) + s_i, \quad i = 1, 2, \cdots, n \)

Output equation: \( y_i(t) = \sigma_i(x_i(t)), \quad i = 1, 2, \cdots, n \) (1.1)

where \( x_i \) represents the state of the \( i \)th neuron, \( y_i \) is the output of the \( i \)th neuron, \( w_{ij} \) is the synaptic connection weight from the \( i \)th neuron to the \( j \)th neuron, \( s_i \) is a constant external input, \( \sigma_i(\cdot) \) is the activation function. This nonlinear system can be implemented by an analog RC (resistance-capacitance) network circuit is shown in Figure 1.4. A circuit contains a RC network at the input of each amplifier. The capacitance \( C_i > 0 \) and the resistance \( R_i > 0 \) represent the total shunt capacitance and shunt resistance at the input of the \( i \)th amplifier. Since the intrinsic delay exhibited by any physical amplifier is modelled by an input resistance \( R_i > 0 \) and \( C_i > 0 \), which are drawn as external components, an actual operational amplifier can, therefore, be assumed as an ideal amplifier without delay. Furthermore, let \( R_{ij} \) be the resistor connecting the output of the \( j \)th amplifier to the input of the \( i \)th amplifier and \( s_i \) the
fixed external input current.

\[ u_i = x_i : \text{input voltage of the } i\text{th amplifier} \]

\[ V_i = \sigma_i(u_i) : \text{output of the } i\text{th amplifier, where each operational amplifier has two output terminals each providing } V_i \text{ and } -V_i. \]

An electronic circuit consisting of operational amplifiers, capacitors and resistors should be able to operate as a Hopfield network. This circuit can be designed by reconstructing the stable states that have been designed using the proper value of \( w_{ij} \) and as long as \( w_{ij} \) is symmetric; that is \( w_{ij} = w_{ji} \) and the amplifiers are quick compared with the characteristic of the neural time constant \( R_iC_i \). In this case, the neural system converges to stable states and will not oscillate or display chaotic behavior. The innovative concepts and implementations of a single-chip electronic NN along the lines just discussed have been reported by several groups using very-large-scale integration.
1.5 Stochastic process

The word “stochastic” means “pertaining to chance” (Greek roots), and is thus used to describe subjects that contain some element of random or stochastic behavior. For a system to be stochastic, one or more parts of the system has randomness associated with it. Unlike a deterministic system, for example, a stochastic system does not always produce the same output for a given input. A few components of systems that can be stochastic in nature include stochastic inputs, random time-delays, noisy (modelled as random) disturbances, and even stochastic dynamic processes.

A stochastic process is one whose behavior is non-deterministic, in which a system’s subsequent state is determined both by the process’s predictable actions and by a random element.

An example of a stochastic process in the natural world is pressure in a gas as modelled by the Wiener process. Even though each molecule is moving in a deterministic path, the motion of a collection of them is computationally and practically unpredictable. A large enough set of molecules will exhibit stochastic characteristics, such as filling the container, exerting equal pressure, diffusing along concentration gradients, etc.

1.5.1 Markov property

In probability theory and statistics, the term Markov property refers to a property of a stochastic process which is initiated by the Russian mathematician Andrey Markov. Moreover, stochastic process has the Markov property if the conditional probability distribution of future states of the process depend only upon the present state; that is, given the present, the future does not depend on the past. A process with this property is called Markov process. The term strong Markov property is similar to this, except that the meaning of “present” is defined in terms of a certain type of random
variable, which might be specified in terms of the outcomes of the stochastic process itself, known as stopping time.

An extension of the Markov property to other circumstances is encompassed by the idea of a Markov random field, which extends the formulation of the property to apply to two or more dimensions, or to random variables defined for an interconnected network of items.

### 1.5.2 Markov jump process

One of the main issues in control systems is their capability of maintaining an acceptable behavior and meeting some performance requirements even in the presence of abrupt changes in the system dynamics. These changes can be due to abrupt environmental disturbances, component failures or repairs, changes in subsystems interconnections, abrupt changes in the operation point for a non-linear plant, etc. Examples of these situations can be found, in economic systems, aircraft control systems, control of solar thermal central receivers, robotic manipulator systems, large flexible structures for space stations, etc. In some cases these systems can be modelled by a set of discrete-time linear systems with modal transition given by a Markov chain. This family is known in the specialized literature as Markov jump linear systems.

MJPs play an important role in a large number of application domains. However, realistic systems are analytically intractable and they traditionally have been analyzed using simulation based techniques, which do not provide a framework for statistical inference. A mean field approximation is proposed to perform posterior inference and parameter estimation. The approximation allows a practical solution to the inference problem, while still retaining a good degree of accuracy. We illustrate our approach on two biologically motivated systems. MJPs provide a rigorous probabilistic framework to model the joint dynamics of groups (species) of interacting individuals, with applications ranging from information packets in a telecommunications network to
epidemiology and population levels in the environment. These processes are usually non-linear and highly coupled, giving rise to non-trivial steady states (often referred to as emerging properties). Unfortunately, this also means that exact statistical inference is unfeasible and approximations must be made in the analysis of these systems. A traditional approach, which has been very successful throughout the past century, is to ignore the discrete nature of the processes and to approximate the stochastic process with a deterministic process whose behavior is described by a system of non-linear and coupled ODEs.

1.5.3 Brownian motion

Brownian motion named after the Scottish Botanist Robert Brown is the seemingly random movement of particles suspended in a fluid (i.e. a liquid such as water or air) or the mathematical model used to describe such random movements, often called a particle theory. Brown in 1826 - 27 observed the irregular motion of pollen particles suspended in water. He and others noted that

♦ the path of a given particle is very irregular, having a tangent at no point, and
♦ the motions of two distinct particles appear to be independent.

The mathematical model of Brownian motion has several real-world applications including stock market fluctuations. However, movements in share prices may arise due to unforeseen events which do not repeat themselves, further physical and economic phenomena are not comparable. Brownian motion is among the simplest of the continuous-time stochastic processes, and it is a limit of both simpler and more complicated stochastic processes. This universality is closely related to that of the normal distribution. In both cases, it is often mathematical convenience rather than the accuracy of the models that motivates their use.

In 1900, first time Bachelier attempted to describe fluctuations in stock prices mathematically and essentially discovered certain results later rederived and extended
by Einstein in 1905 (for which he received the Nobel prize in 1921). The motion
was later explained by random collisions with the molecules of water. To describe it
mathematically, we use the concept of a stochastic process \( B_t(\omega) \), interpreted as the
position of the pollen grain \( \omega \) at time \( t \).

For a dynamic system the simplest continuous stochastic perturbation is naturally
considered to be a Brownian motion, since it is a Normal process (or say, a Guassian
process) with independent increments which are also normally distributed. In general,
a continuous stochastic perturbation will be modelled as some stochastic integral with
respect to the Brownian motion.

1.6 Stochastic differential equations

SDE is a combination of DEs, probability theory and stochastic process. SDEs arise in
modeling of variety of random dynamic phenomena in the physical, biological, Engi-
neering and social sciences. In some real physical Engineering problems such as wind
excitation or seismic impact, it is very difficult to describe the dynamic behavior of
the system by a mathematical model. The possible way to model these excitations is
by the use of probabilistic concepts instead of deterministic one. In particular, SDEs
are used with increasing frequency in a diverse range of fields. Financial engineers
use SDEs as the basis of stochastic volatility models. These are used for modeling
neurons in computational cell biology, model of Brownian motion in Physics, study
of single molecule fluorescence and also have a potential application as computational
tools for Chemistry. Also they are used in the study of Seismology, Hydrology and
fatigue testing in Engineering.

In sciences, SDEs are used to model systems that are inherently random or subject
to random external perturbations. Furthermore, systems in continuum mechanics or in
financial economics have governing equations, which involve integral terms representing
the effect of the past. All kinds of dynamics with stochastic influence in nature or complex system created by mankind are modelled by SDEs.

(i) Seed dispersal model

Considering the estimation of the horizontal distance that certain seeds traverse when falling from a given height under the influence of a randomly varying wind. It is assumed that the seeds experience a frictional force proportional to the square of the speed of the air on the seeds.

First, consider a deterministic model where the wind speed is constant and not varying randomly. Let $V_\omega(t) = v_\omega(t)i$ be the wind velocity, $V(t) = v_x(t)i + v_y(t)j$ be the seed velocity, and $V_a(t) = V_\omega(t) - V(t)$ be the air velocity on the seed at time $t$. Let $F_f(t) = k|V_a(t)|^2 \frac{V_a(t)}{|V_a(t)|} = k|V_a(t)||V_a(t)$ be the frictional force on the seed at time $t$, where $k$ is a constant of proportionality. Let $F_g(t) = -mgj$ be the force of gravity on the seed of mass $m$. Finally, let $s(t) = x(t)i + y(t)j$ be the position of the seed at time $t$. It is assumed that $V(0) = 0 = 0i + 0j$ and $s(0) = hj$ where $h$ is the initial height. It is straightforward to check that the velocity and the position of the seed at any time $t$ satisfy the following initial-value system

$$\begin{align*}
\frac{ds(t)}{dt} &= V(t), \\
\frac{dV(t)}{dt} &= \frac{1}{m} \left( F_f(t) + F_g(t) \right) = \frac{k}{m} |V_a(t)||V_a - gj, \\
V_a(t) &= V_m(t) - V(t) \text{ with } V_\omega(t) = v_\omega i, \\
s(0) &= hj, \quad V(0) = 0.
\end{align*}$$

(1.2)

Now suppose that the wind speed randomly varies about a mean speed $v_\omega$. In particular, the wind speed can experience a change of $\pm \alpha$ during a small time interval $\Delta t$ with the probabilities listed in Table 1.1 here it is assumed that $\Delta t$ is sufficiently small so that $p_1, p_2 > 0$ (Notice that $p_1$ and $p_2$ can be defined as $p_l = \max\left\{0, \left(\lambda + (-1)^l \beta(v_v - v_\omega(t))\right)\Delta t\right\}$ for $l = 1, 2$ to guarantee nonnegativity of the probabilities). The value $\lambda \Delta t$ represents the probability associated with random
diffusion of the wind speed and does not depend on \( v_\omega(t) \). The term \( \pm \beta(v_e - v_\omega(t)) \) represents the probability associated with drift towards the mean wind speed of \( v_e \). When \( v_\omega(t) \neq v_e \), the probability that the wind speed moves closer to \( v_e \) is greater than the probability that the wind speed moves further from \( v_e \). Thus, \( v_e \) can be regarded as a mean wind speed. The next step is to find the mean wind speed change and the variance in the change. It is straightforward to show that the required expectations to order \((\Delta t)^2\) are

\[
E(\Delta v_\omega) = 2\alpha \beta (v_e - v_\omega(t)) \Delta t
\]

and

\[
E((\Delta v_\omega)^2) = 2\alpha^2 \lambda \Delta t.
\]

Based on the above arguments, a reasonable SDE model for the randomly wind speed is

\[
dv_\omega(t) = 2\alpha \beta (v_e - v_\omega(t))dt + \sqrt{2\alpha^2 \lambda}dW(t).
\]

Indeed, this SDE can be solved exactly to yield

\[
v_\omega(t) = v_e + \exp(-2\alpha \beta t) \left( - v_e + v_\omega(0) + \int_0^t \sqrt{2\alpha^2 \lambda} \exp(2\alpha \beta s)dW(s) \right).
\]

This solution implies, for large time \( t \), that the wind speed \( v_\omega(t) \) is approximately normally distributed with mean \( v_e \) and variance \( 2\alpha^2 \lambda/(4\alpha \beta) = \alpha \lambda/(2\beta) \).
The complete model for the seed dispersal dynamics in a randomly varying wind is now formulated as

\[
\begin{align*}
\frac{ds(t)}{dt} &= V(t), \\
\frac{dV(t)}{dt} &= \frac{1}{m} \left( F_f(t) + F_g(t) \right) = \frac{k}{m} |V_a(t)| V_a - g_f, \\
\frac{d\omega(t)}{dt} &= 2\alpha\beta(\nu_e - \nu_\omega(t)) + \sqrt{2\alpha^2\lambda} \frac{dV(t)}{dt}, \\
V_a(t) &= V_\omega(t) - V(t), \\
s(0) &= h_j, \quad V(0) = 0, \quad V_\omega(0) = \nu_\omega^0.
\end{align*}
\]

This problem illustrates the ease with which a deterministic model can be transformed into a SDE model for certain physical problems after agreeing upon a discrete stochastic model. Notice that there are other interesting possibilities for modeling the wind speed besides the model described here. For example, in constructing the probabilities for the wind speed changes, consider \( p_l = (\lambda + (-1)^l \beta(\nu_e - \nu_\omega(t)))^{k_1} \Delta t \) for \( l = 1, 2 \) for some positive odd integers \( k_1 \) and \( k_2 \). Selecting the parameter values and the discrete stochastic process that represents a given physical situation may involve many computational comparisons with physical data.

### 1.7 Stability

In 1892, Lyapunov introduced the concept of stability of a dynamic system. Roughly speaking, the stability means insensitivity of the state of the system to small changes in the initial state or the parameters of the system. For a stable system, the trajectories which are “close” to each other at a specific instant should therefore remain close to each other at all subsequent instants. Many parts of the qualitative theory of differential equations and dynamical systems deal with asymptotic properties of solutions and the trajectories. The simplest kind of behavior is exhibited by equilibrium points, or fixed points, and by periodic orbits. If a particular orbit is well understood, it is natural
to ask whether a small change in the initial condition will lead to similar behavior. Stability means that the trajectories do not change too much under small perturbations. The opposite situation, where a nearby orbit is getting repelled from the given orbit, is also of interest. In general, perturbing the initial state in some directions results in the trajectory asymptotically approaching the given one and in other directions to the trajectory getting away from it. There may also be directions for which the behavior of the perturbed orbit is more complicated (neither converging nor escaping completely), and then stability theory does not give sufficient information about the dynamics. One of the key ideas in stability theory is that the qualitative behavior of an orbit under perturbations can be analyzed using the linearization of the system near the orbit. In particular, at each equilibrium of a smooth dynamical system with an $n$-dimensional phase space, there is a certain $n \times n$ matrix $A$ whose eigenvalues characterize the behavior of the nearby points. More precisely, if all eigenvalues are negative real numbers or complex numbers with negative real parts then the point is a stable attracting fixed point and the nearby points converge to it at an exponential rate. If none of the eigenvalues is purely imaginary (or zero) then the attracting and repelling directions are related to the eigen spaces of the matrix $A$ with eigenvalues whose real part is negative and positive respectively. Analogous statements are known for perturbations of more complicated orbits.

Concept of stability in common and Engineering sense reflects necessity to keep response of a disturbed system within accepted limits. If deviations describing response of the system from a given regime (e.g. state of equilibrium) lie within prescribed limits, the system is called stable. Otherwise, the system is called unstable. Disturbances, response and prescribed limits can be specified in each case in different ways.
1.8 Lyapunov stability theory

In Mathematics, stability theory addresses the stability of solutions of differential equations and of trajectories of dynamical systems under small perturbations of initial conditions. Lyapunov developed the stability theory of dynamical systems determined by nonlinear time-varying ODEs. In the concepts of stability and instability, he has developed two general methods for the stability analysis of an equilibrium: Lyapunov’s direct method, also called the second method of Lyapunov, and the indirect method of Lyapunov, also called the first method. The former involves the existence of scalar-valued auxiliary functions of the state space (called Lyapunov functions) to ascertain the stability properties of an equilibrium, whereas the latter seeks to deduce the stability properties of an equilibrium of a system described by a nonlinear DE from the stability properties of its linearization. In the process of discovering the first method, Lyapunov established some important stability results for linear systems (involving the Lyapunov matrix equation).

An orbit is called Lyapunov stable if the forward orbit of any point in a small enough neighborhood of it stays in a small (but perhaps, larger) neighborhood. Various criteria have been developed to prove stability or instability of an orbit. Under favorable circumstances, the problem of stability may be reduced to a well-studied problem involving eigenvalues of matrices. A more general method involves Lyapunov functions where as Lyapunov stability theorems give only sufficient condition.

Lyapunov method is a powerful method for determining the stability or instability of fixed points of nonlinear autonomous systems. In mathematics, Lyapunov functions are functions which can be used to prove the stability of a certain fixed point in a dynamical system or autonomous DE.

One must be aware that the basic Lyapunov theorems for autonomous systems are sufficient, but not necessary tool to prove the stability of an equilibrium. Finding a Lyapunov function for a certain equilibrium might be a matter of luck. Trial and
error is the method to apply, when testing Lyapunov-candidate functions on some equilibrium.

1.9 Basic theorems of Lyapunov

Lyapunov theorem states as follows: Let $V(x, t)$ be a non-negative function with derivative $\dot{V}(x, t)$ along the trajectories of the system.

✗ If $V(x, t)$ is locally positive definite and $\dot{V}(x, t) \leq 0$ locally in $x$ and for all $t$, then the origin of the system is locally stable (in the sense of Lyapunov).

✗ If $V(x, t)$ is locally positive definite and decrescent, further $\dot{V}(x, t) \leq 0$ locally in $x$ and for all $t$, then the origin of the system is uniformly locally stable (in the sense of Lyapunov).

✗ If $V(x, t)$ is locally positive definite and decrescent, further $-\dot{V}(x, t)$ is locally positive definite, then the origin of the system is uniformly locally asymptotically stable.

✗ If $V(x, t)$ is positive definite and decrescent, also $-\dot{V}(x, t)$ is positive definite, then the origin of the system is globally uniformly asymptotically stable.

The above theorem gives sufficient conditions for the stability of the origin of a system. It does not, however, give a prescription for determining the Lyapunov function $V(x, t)$. Since the theorem only gives sufficient conditions, the search for a Lyapunov function establishing stability of an equilibrium point could be arduous. However, it is a remarkable fact that the converse of theorem also exists: if an equilibrium point is stable, then there exists a function $V(x, t)$ satisfying the conditions of the theorem. However, the utility of this and other converse theorems is limited by the lack of a computable technique for generating Lyapunov functions. Theorem also stops short of
giving explicit rates of convergence of solutions to the equilibrium. It may be modified to do so in the case of exponentially stable equilibria.

1.10 Basic ideas for Lyapunov’s framework

In the Lyapunov framework, there are two basic ideas

- To construct some simple quadratic Lyapunov candidates (Lyapunov-Krasovskii functional, or Lyapunov-Razumikhin functions) leading to some sufficient stability conditions, more or less conservative.

- To construct more “complicated” quadratic Lyapunov candidates, which lead to necessary and sufficient asymptotic stability conditions, but which are difficult to check for practical problem. One of the ideas to handle such situation is to use discretization techniques.

In both cases, the corresponding stability condition will be expressed in terms of LMIs, as feasibility (delay-independent stability) or optimization problems (computing the maximal allowable delay, or maximal allowable ellipsoids in the delay-parameter space, etc).

1.11 A brief history of LMIs in control theory

The history of LMIs in the analysis of dynamical systems goes back more than 100 years. The story begins in about 1890, when Lyapunov published his seminal work introducing what we now call Lyapunov theory. He showed that the DE

\[
\frac{d}{dt} x(t) = Ax(t)
\]  

(1.3)
is stable (i.e., all trajectories converge to zero) if and only if there exists a positive-definite matrix $P$ such that

$$A^TP + PA < 0.$$  \hspace{1cm} (1.4)

The requirement $P > 0$, $A^TP + PA < 0$ is what we now call a Lyapunov inequality on $P$ which is a special form of an LMI. Lyapunov also showed that this first LMI could be explicitly solved. Indeed, we can pick any $Q = Q^T > 0$ and then solve the linear equation $A^TP + PA = -Q$ for the matrix $P$, which is guaranteed to be positive-definite if the system (1.3) is stable. In summary, the first LMI used to analyze the stability of a dynamical system was the Lyapunov inequality (1.4), which can be solved analytically (by solving a set of linear equations). The next major milestone occurs in 1940’s when Lur’e, Postnikov and others in the Soviet Union applied Lyapunov’s methods to some specific practical problems in control Engineering, especially, the problem of stability of a control system with a nonlinearity in the actuator. Although they did not explicitly form matrix inequalities, their stability criteria have the form of LMIs. These inequalities were reduced to polynomial inequalities which can be checked easily “by hand” (for, needless to say, small systems). Nevertheless they were justifiably excited by the idea that Lyapunov’s theory could be applied to important (and difficult) practical problems in control Engineering. A summary of key events in the history of LMIs in control theory are given as follows:

- 1890: First LMI appears; analytic solution of the Lyapunov LMI via Lyapunov equation.

- 1940’s: Application of Lyapunov’s methods to real control Engineering problems. Small LMIs have been solved “by hand”.

- Early 1960’s: Positive-real lemma gives graphical techniques for solving another family of LMIs.

- Late 1960’s: Observation that the same family of LMIs can be solved by solving an ARE.
• Early 1980's: Recognition that many LMIs can be solved by computer via convex programming.

• Late 1980's: Development of interior-point algorithms for LMIs.

1.12 Linear matrix inequalities

LMIs and LMI techniques have emerged as powerful design tools in areas ranging from control Engineering to system identification and structural design. Three factors make LMI techniques appealing:

❖ A variety of design specifications and constraints can be expressed as LMIs.

❖ Once formulated in terms of LMIs, a problem can be solved exactly by efficient convex optimization algorithms (the "LMI solvers").

❖ While most problems with multiple constraints or objectives lack analytical solutions in terms of matrix equations, they often remain tractable in the LMI framework. This makes LMI-based design a valuable alternative to classical "analytical" methods.

LMI has the form

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0,$$  \hspace{1cm} (1.5)

where \(x \in \mathbb{R}^m\) is the variable and the symmetric matrices \(F_i = F_i^T \in \mathbb{R}^{n \times n}\), \(i = 0, \ldots, m\) are given. The inequality symbol in (1.5) means that \(F(x)\) is positive-definite, that is, \(u^T F(x) u > 0\) for all nonzero \(u \in \mathbb{R}^n\). Of course, the LMI (1.5) is equivalent to a set of \(n\) polynomial inequalities in \(x\), that is, the leading principal minors of \(F(x)\) must be positive.
The LMI (1.5) is a convex constraint on $x$, that is, the set \( \{ x \mid F(x) > 0 \} \) is convex. Although the LMI (1.5) may seem to have a specialized form, it can represent a wide variety of convex constraints on $x$. In particular, linear inequalities, (convex) quadratic inequalities, matrix norm inequalities, and constraint that arise in control theory, such as Lyapunov and convex quadratic matrix inequalities can all be cast in the form of LMI.

Multiple LMIs $F^{(1)}(x) > 0$, $\ldots$, $F^{(p)}(x) > 0$ can be expressed as the single LMI given as $\text{diag}(F^{(1)}(x), \ldots, F^{(p)}(x)) > 0$. Therefore, we will make no distinction between a set of LMIs and a single LMI.

1.13 Stability of time-delay systems

The stability of time-delay systems has been widely investigated in the last two decades. Practical examples of time-delay systems include chemical Engineering, communications and biological systems. Current efforts can be divided into two classes: namely, frequency-domain approach and time-domain approach. In the time-domain approach, the direct Lyapunov method is a powerful tool. There are two different ideas how one can use this method. They are the Lyapunov-Krasovskii approach and the Lyapunov-Razumikhin approach. In the frequency domain approach, two variables polynomial method, matrix pencils, $\mu$-analysis, IQC analysis are considered in the case. Both approaches can be used to handle systems with time-varying delay. The former usually requires both the upper bound of the time-varying delay and additional information on the derivative of the time-varying delay, while the latter has no restriction on the derivative of the time-varying delay, which allows a fast time-varying delay. The obtained results using the Lyapunov-Krasovskii approach are usually less conservative than those using the Lyapunov-Razumikhin approach since the former takes advantage of the additional information of the delay. It is well known that there are systems which are stable with some nonzero delay, but are unstable without delay. For such
case, if there is a time-varying perturbation on the nonzero delay, it is of great significance to consider the stability of systems with interval time-varying delay. Other typical examples of systems with interval time-varying delay are networked control systems. Time-varying delay is relevant when considering the followings

- Network congestion control - delay depends on the length of queue and network traffic
- Real-time control systems with dynamic scheduling - delay depends on the size of the processes
- Control of chemical processes
- Biological systems

Two different types of stability criteria can be considered

- Delay-independent stability - stability for any length of delay.
- Delay-dependent stability - assume a priori knowledge on the upper-bounds of the delay.

One of the most widely used tools for investigating the stability of linear systems is the second (direct) method of Lyapunov, presented in his dissertation of 1892. The idea of this method is to investigate stability of a given system by measuring the rate of change of the energy of the system. The advantage of this approach is that it allows one to infer the stability of differential (and difference) equations without explicit knowledge of solutions.
1.14 Stability of neural networks

Complex nonlinear structures of dynamic NNs used in computing tasks such as information processing, or in associative memory for storing patterns, present a challenge in stability investigations. The notion of the stability of an equilibrium point of a dynamic system is of fundamental importance in dynamic NNs. The stability of the equilibrium points of a dynamic NN is one of the most basic and important properties for many Engineering applications. For dynamic NNs, we refer to stability in the sense of Lyapunov. The dynamic behavior and the notions of stability of continuous time-dynamic NNs described by a set of nonlinear differential equations have been widely studied since the early 1990s.

As mentioned before, time-delay appears in the electronic implementations of NNs and can lead to complicated dynamic behaviors such as chaos, oscillation and instability. In fact, the NNs applied to optimization problems in [93] must have a unique and stable equilibrium point. That is, stability is one of the major properties of NNs and is a crucial feature in the design of NNs. The stability analysis of NNs with time-delays has thus received a great deal of attention. Various approaches, including the nonsingular $M$-matrix based approach [5], nonlinear measure approach, nonsmooth analysis approach [18] as well as LMI approach, have been developed for the stability analysis of DNNs. In [107], the global asymptotic stability of DNNs is discussed, where the activation functions may be non-Lipschitz continuous. Some inequalities are exploited to obtain more general stability criteria, which can be applied to a broad range of activation functions.

Most of the results on stability analysis of DNNs built on the above-mentioned approaches are independent of time-delays. In other words, the information of time-delays has not been taken into account when deriving stability conditions. Therefore, they tend to be conservative in general. Recently, the LMI technique is fruitfully applied to investigate the stability of DNNs. The main advantage of this approach
is that delay-dependent, easily-verified and less conservative stability results can be established. The approaches of stability analysis of time-delay linear systems stated in previous section can be efficiently borrowed to discuss the stability of DNNs. For example, the free-weighting matrix approach has been applied in [64, 94]. Many other stability results of DNNs have also been reported in the literature see, for example [3, 11, 56, 105] and references therein.

Since NNs usually have a spatial extend due to the presence of a multitude of parallel pathways with a variety of axon sizes and length. Hence there is a distribution of propagation delays over a period of time. It is worth noting that, although the signal propagation is sometimes instantaneous and can be modelled with discrete delays, it may also be distributed during a certain time period so that the distributed delays should be incorporated in the model. In other words, it is often the case that the NN model possesses both discrete and distributed delays [9, 13, 58]. Recently, it is noted that stability analysis of HNNs, CGNNs and BAMNNs with distributed delays have been discussed in [32, 86] and the references therein. In addition, in hardware implementation of NNs, stochastic disturbances are nearly inevitable owing to thermal noise in electronic devices. Due to stochastic disturbances, stability of the NNs may be affected. Recently, some results on stability of SNNs with time-varying delays have been reported in [38, 39, 41, 83, 89, 106] and the references therein.

1.15 Robust stability analysis

Robustness is an approach to feature persistence in systems for which we do not have the mathematical tools to use the approaches of stability theory. In some cases the problem could be reformulated as one of stability theory, but only in a formal sense that would bring little in the way of new insight or control methodologies. The robust stability problem considers the stability problem of systems that contain some uncertainties.
In Engineering applications, it is now very common that one does not know exactly the system under investigation; that is, the system contains some elements (blocks) that are uncertain. Usually it is known that these uncertain elements belong to some specific admissible domains, which in turn depend on the nature of the elements and also on the information available about the system. In other words, it is known only that the system belongs to the family of systems that arises when the uncertain elements (blocks) range over the admissible domains and therefore one may treat the family as a new object for analysis. This family is referred to as an uncertain system. When it is possible to show that all systems of the family are stable, the stability of the original system that is a particular member of the family is guaranteed.

As is well known, it is usually impossible to describe a practical system exactly. First, there are often parameters or parasitic processes that are not completely known. Second, due to the limitation of mathematical tools available, we usually try to use a relatively simple model to approximate a practical system. As a result, some aspects of the system dynamics (known as unmodeled dynamics) are ignored. Third, some control systems are required to operate within a range of different operating condition. To capture these uncertain factors, it is often possible to identify a bounding set such that all the possible uncertainties fall within this set and yet it, is not too difficult to analyze mathematically.

1.15.1 Uncertainty characterization

Consider the system

$$\dot{x}(t) = A_0(x_t, t)x(t) + A_1(x_t, t)x(t - r), \quad (1.6)$$

where $A_0 \in \mathbb{R}^{n \times n}, A_1 \in \mathbb{R}^{n \times n}$ are uncertain coefficient matrices not known completely, except that they are within a compact set $\mathcal{U}$ which refer the uncertainty set

$$(A_0(x_t, t), A_1(x_t, t)) \in \omega \text{ for all } t \geq 0.$$
The uncertainty set characterizes the uncertainties and serves as basic information needed to carry out robust stability analysis. Notice also that the coefficients may depend on the time \( t \) as well as the current and previous state variable \( x(t+\xi), -r \leq \xi \leq 0 \). For the sake of convenience, we will not explicitly show these dependences or only show the dependence on time \( t \) when no confusion may arise.

A good choice of uncertainty set is a compromise between minimizing conservatism (and therefore, it is desirable to make the uncertainty set "small") and the mathematical tractibility (and therefore, it is desirable to make the uncertainty set structurally simple).

1.15.2 Polytopic uncertainty

The first class of uncertainty frequently encountered in practice is the polytopic uncertainty. In this case, there exist, say, \( n_\nu \) elements of the uncertainty set \( \tilde{U} \)

\[
\tilde{\omega}^{(k)} = \left( A_0^{(k)}, A_1^{(k)} \right) \quad k = 1, 2, \ldots, n_\nu
\]

known as vertices, such that \( \tilde{U} \) can be expressed as the convex hull of these vertices

\[
\tilde{U} = \text{co}\{\tilde{\omega}^{(k)} \mid k = 1, 2, \ldots, n_\nu \}.
\]

In other words, the uncertainty set \( \tilde{U} \) consists of all the convex linear combination of the vertices

\[
\tilde{U} = \left\{ \sum_{k=1}^{n_\nu} \alpha_k \tilde{\omega}^{(k)} \mid \alpha_k \geq 0, k = 1, 2, \ldots, n_\nu; \sum_{k=1}^{n_\nu} \alpha_k = 1 \right\}.
\]

For example, in practice, there are often some uncertain parameters in the system. Each uncertain parameter may vary between a lower limit and an upper limit, and these uncertain parameters often appear linearly in the system matrices. In this case, the collection of all the possible system matrices form a polytopic set. The vertices of this set can be calculated by setting the parameters to either lower or upper limit. If
there are $n_p$ uncertain parameters, it is easy to see that there are $n_v = 2^n$ uncertain parameters, it is easy to see that there are two uncertain parameters, $\alpha$ and $\beta$, varying between lower and upper bounds

$$\alpha_{\min} \leq \alpha \leq \alpha_{\max}, \quad \beta_{\min} \leq \beta \leq \beta_{\max}.$$  

They appear linearly in the system matrices

$$A_0 = A_{0n} + \alpha_0 \alpha + \beta A_{0\beta},$$

$$A_1 = A_{1n} + \alpha_1 \alpha + \beta A_{1\beta},$$

or in the abbreviated form

$$\bar{\omega} = \bar{\omega}_n + \alpha \bar{\omega}_\alpha + \beta \bar{\omega}_\beta.$$

Then, the uncertainty set possess four vertices.

$$\bar{\omega}^{(1)} = \bar{\omega}_n + \alpha_{\min} \bar{\omega}_\alpha + \beta_{\min} \bar{\omega}_\beta$$

$$\bar{\omega}^{(2)} = \bar{\omega}_n + \alpha_{\max} \bar{\omega}_\alpha + \beta_{\min} \bar{\omega}_\beta$$

$$\bar{\omega}^{(3)} = \bar{\omega}_n + \alpha_{\min} \bar{\omega}_\alpha + \beta_{\max} \bar{\omega}_\beta$$

$$\bar{\omega}^{(4)} = \bar{\omega}_n + \alpha_{\max} \bar{\omega}_\alpha + \beta_{\max} \bar{\omega}_\beta.$$

### 1.15.3 Norm bounded uncertainty

In terms of norm-bounded uncertainty, we decompose the system matrices $\bar{\omega} = (A_0, A_1)$ into two parts: the nominal part $\bar{\omega}_n = (A_{0n}, A_{1n})$ and the uncertain part $\Delta \bar{\omega} = (\Delta A_0, \Delta A_1)$:

$$\bar{\omega} = \bar{\omega}_n + \Delta \bar{\omega},$$

that is,

$$A_0 = A_{0n} + \Delta A_0,$$

$$A_1 = A_{1n} + \Delta A_1.$$
The uncertain part is written as
\[
(\Delta A_0 \quad \Delta A_1) = EF(G_0 \quad G_1),
\]
where \( E, G_0, \) and \( G_1 \) are known constant matrices. \( F \) is an uncertain matrix satisfying
\[
\|F\| \leq 1.
\]
In other words, the uncertainty set \( \mathcal{U} \) can be expressed as
\[
\mathcal{U} = \{(A_{0n} + EFG_0, A_{1n} + EFG_1) \mid \|F\| \leq 1\}.
\]

A slightly more general uncertainty is the linear fractional norm-bounded uncertainty (or LF norm-bounded uncertainty), in which the uncertain matrix \( F \) is replaced by
\[
(\Delta A_0 \quad \Delta A_1) = E(I - FD)^{-1}F(G_0 \quad G_1)
\]
and satisfies \( \|F\| \leq 1 \). Clearly, the norm-bounded uncertainty is a special case of LF norm-bounded uncertainty when \( D = 0 \). Of course, for the LF norm-bounded uncertainty to be well defined, the matrix \( I - FD \) has to be invertible for arbitrary \( F \) satisfying \( \|F\| \leq 1 \). It is not difficult to see that this is equivalent to requiring
\[
I - D^TD > 0.
\]

1.16 Basic definitions and useful lemmas

To end this chapter, some useful definitions and lemmas are stated below which are useful to derive the main results in the following chapters.

**Definition 1.16.1** All the possible outcomes, the elementary events are grouped together to form a set \( \Omega \) with typical element \( \omega \in \Omega \). Not every subset of \( \Omega \) is in general an observable or interesting event. So we only group these observable or interesting events together as a family \( \mathcal{F} \) of subsets of \( \Omega \). For the purpose of probability theory, such a family \( \mathcal{F} \) should have the following properties:
(i) $\emptyset \in \mathcal{F}$, where $\emptyset$ denotes the empty set,

(ii) $\alpha \in \mathcal{F} \Rightarrow \alpha^C \in \mathcal{F}$, where $\alpha^C = \Omega - \alpha$ is the complement of $\alpha$ in $\Omega$,

(iii) $\{\alpha_i\}_{i \geq 1} \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} \alpha_i \in \mathcal{F}$.

A family $\mathcal{F}$ with these three properties is called a $\sigma$-algebra. The pair $(\Omega, \mathcal{F})$ is called a measurable space, and elements of $\mathcal{F}$ is henceforth called $\mathcal{F}$-measurable sets instead of events.

**Definition 1.16.2** A real valued function $x : \Omega \to \mathbb{R}$ is said to be $\mathcal{F}$-measurable if

$$\{\omega : x(\omega) \leq a\} \text{ for all } a \in \mathbb{R}.$$  

The function $x$ is also called a real-valued ($\mathcal{F}$-measurable) random variable.

**Definition 1.16.3** A probability measure $\mathcal{P}$ on a measurable space $(\Omega, \mathcal{F})$ is a function $\mathcal{P} : \mathcal{F} \to [0, 1]$ such that

(i) $\mathcal{P}(\Omega) = 1$;

(ii) for any disjoint sequence $\{\alpha_i\}_{i \geq 1} \subset \mathcal{F}$ (that is $\alpha_i \cap \alpha_j = \emptyset$ if $i \neq j$)

$$\mathcal{P}
\left(\bigcup_{i=1}^{\infty} \alpha_i\right) = \sum_{i=1}^{\infty} \mathcal{P}(\alpha_i).$$

The triple $(\Omega, \mathcal{F}, \mathcal{P})$ is called a probability space.

**Definition 1.16.4** If $(\Omega, \mathcal{F}, \mathcal{P})$ is a probability space, we set

$$\overline{\mathcal{F}} = \left\{ \alpha \subset \Omega : \exists \beta, \gamma \in \mathcal{F} \text{ such that } \beta \subset \alpha \subset \gamma, \quad \mathcal{P}(\beta) = \mathcal{P}(\gamma) \right\}.$$  

Then $\overline{\mathcal{F}}$ is a $\sigma$-algebra and is called the completion of $\mathcal{F}$. If $\mathcal{F} = \overline{\mathcal{F}}$, the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ is said to be complete.

**Definition 1.16.5** (Stochastic processes) Families of random variables which are functions of time, are known as stochastic processes (or random processes or random functions).
Let \((\Omega, \mathcal{F}, \mathcal{P})\) be a probability space. A **filtration** is a family \(\{\mathcal{F}_t\}_{t \geq 0}\) of increasing sub \(\sigma\)-algebras of \(\mathcal{F}\) (that is \(\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}\) for all \(0 \leq t < s < \infty\)). The filtration is said to be **right continuous** if \(\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s\) for all \(t \geq 0\). When the probability space is complete, the filtration is said to satisfy the usual conditions if it is right continuous and \(\mathcal{F}_0\) contains all \(\mathcal{P}\)-null sets.

From now on, unless otherwise specified, we shall always work on a given complete probability space \((\Omega, \mathcal{F}, \mathcal{P})\) with a filtration \(\{\mathcal{F}_t\}_{t \geq 0}\) satisfying the usual conditions.

A family \(\{x(t)\}_{t \in I}\) of \(\mathbb{R}^d\)-valued random variables is called a **stochastic process** with parameter set \((\text{or index set})\) \(I\) and state space \(\mathbb{R}^d\). The parameter set \(I\) is usually the halfline \(\mathbb{R}_+ = [0, \infty)\), but it may also be an interval \([a, b]\), the nonnegative integers or even subsets of \(\mathbb{R}^d\). Note that for each fixed \(t \in I\), we have a random variable
\[
\omega \rightarrow x(t, \omega) \in \mathbb{R}^d \text{ for every } \omega \in \Omega.
\]

On the other hand, for each fixed \(\omega \in \Omega\), we have a function
\[
t \rightarrow x(t, \omega) \in \mathbb{R}^d \text{ for every } t \in I
\]
which is called a sample path of the process.

Let \(\{x(t)\}_{t \geq 0}\) be an \(\mathbb{R}^d\)-valued stochastic process. It is said to be \(\{\mathcal{F}_t\}\)-**adapted** (or simply, **adapted**) if for every \(t\), \(x(t)\) is \(\mathcal{F}_t\)-measurable.

**Definition 1.16.6** Let \((\Omega, \mathcal{F}, \mathcal{P})\) be a probability space with a filtration \(\{\mathcal{F}_t\}_{t \geq 0}\). A (standard) one-dimensional Brownian motion is a real-valued continuous \(\{\mathcal{F}_t\}\)-adapted process \(\{W(t)\}_{t \geq 0}\) with the following properties:

(i) \(W(0) = 0\) almost surely,

(ii) for \(0 \leq s < t < \infty\), the increment \(W(t) - W(s)\) is normally distributed with mean zero and variance \(t - s\),

(iii) for \(0 \leq s < t < \infty\), the increment \(W(t) - W(s)\) is independent of \(\mathcal{F}_s\).
Definition 1.16.7 The solution \( x = 0 \) of the system described by the equation

\[
\dot{x}(t) = f(t, x(t)), \quad x(t_0) = x_0,
\]

where \( x \) and \( x_0 \) are elements of \( \mathbb{R}^n \) is said to be stable if for every \( \epsilon > 0 \), there exist a \( \delta(t_0, \epsilon) > 0 \) such that \( \|x(t_0)\| < \delta \) implies \( \|x(t)\| < \epsilon \), for every \( t \geq t_0 \).

Definition 1.16.8 The solution \( x = 0 \) of the system (1.7) is said to be asymptotically stable if it is stable and if there exists a \( \delta(t_0) > 0 \) such that \( \|x(t_0)\| < \delta \) implies \( \lim_{t \to \infty} x(t) = 0 \).

Definition 1.16.9

(i) The trivial solution of equation

\[
dx(t) = f(x(t), t)dt + g(x(t), t)dB(t) \quad \text{on} \quad t \geq t_0
\]

is said to be stochastically stable or stable in probability if for every pair of \( \epsilon \in (0, 1) \) and \( r > 0 \), there exists a \( \delta = \delta(\epsilon, r, t_0) > 0 \) such that

\[
P\left\{ |x(t; t_0, x_0)| < r \text{ for all } t \geq t_0 \right\} \geq 1 - \epsilon
\]

whenever \( |x_0| < \delta \).

(ii) The trivial solution is said to be stochastically asymptotically stable if it is stochastically stable and moreover, for every \( \epsilon \in (0, 1) \), there exists a \( \delta_0 = \delta_0(\epsilon, t_0) > 0 \) such that

\[
P\left\{ \lim_{t \to \infty} x(t; t_0, x_0) = 0 \right\} \geq 1 - \epsilon
\]

whenever \( |x_0| < \delta \).

Lemma 1.16.10 (Schur Complement) [8] Given constant matrices \( \Omega_1, \Omega_2 \) and \( \Omega_3 \) with appropriate dimensions, where \( \Omega_1^T = \Omega_1 \) and \( \Omega_2^T = \Omega_2 > 0 \), then

\[
\Omega_1 + \Omega_3^T \Omega_2^{-1} \Omega_3 < 0
\]

32
if and only if
\[
\begin{bmatrix}
\Omega_1 & \Omega_2^T \\
* & -\Omega_2
\end{bmatrix} < 0, \quad \text{or} \quad \begin{bmatrix}
-\Omega_2 & \Omega_3 \\
* & \Omega_4
\end{bmatrix} < 0.
\]

**Lemma 1.16.11** (Jensen’s inequality) [27] For any \( n \times n \) constant matrix \( M > 0 \), any scalars \( a \) and \( b \) with \( a < b \) and a vector function \( x(t) : [a, b] \rightarrow \mathbb{R}^n \) such that integrations concerned are well defined, then the following inequality holds
\[
\left[ \int_a^b x(s)ds \right]^T M \left[ \int_a^b x(s)ds \right] \leq (b-a) \left[ \int_a^b x^T(s)Mx(s)ds \right].
\]

**Lemma 1.16.12** [102] For any vectors \( x, y \in \mathbb{R}^n \), matrices \( A, P, H, N, F(t) \) are real matrices of appropriate dimensions with \( P > 0, \quad F^T(t)F(t) \leq I \) and scalar \( \epsilon > 0 \), the following inequalities hold

(i) \( H F(t)N + N^T F^T(t)H^T \leq \epsilon^{-1} HH^T + \epsilon N^T N \)

(ii) If \( P^{-1} - \epsilon^{-1} HH^T > 0 \), then
\[
(A + HF(t)N)^T P(A + HF(t)N) \leq A^T (P^{-1} - \epsilon^{-1} HH^T)^{-1} A + \epsilon N^T N.
\]

**Lemma 1.16.13** [54] Suppose \( A(t) \) to be given by (3.6)-(3.8). Given matrices \( M = M^T \), \( S \) and \( N \) of appropriate dimensions, the inequality
\[
M + SA(t)N + N^T A^T(t)S^T < 0,
\]
holds for \( F(t) \) such that \( F^T(t)F(t) \leq I \) if and only if for some \( \delta > 0 \)
\[
\begin{bmatrix}
M & S & \delta N^T \\
S^T & -\delta I & \delta J^T \\
\delta N & \delta J & -\delta I
\end{bmatrix} < 0.
\]

**Lemma 1.16.14** For any matrices \( Q_1 = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} > 0 \), \( Q_2 = \begin{bmatrix} \hat{Q}_{11} & \hat{Q}_{12} \\ \hat{Q}_{21} & \hat{Q}_{22} \end{bmatrix} > 0 \), \( M_j, N_j, S_j \), \( (j = 1, 2, \cdots, 20) \) and scalars \( h_1, h_2 \) the following inequalities hold
\[
-\int_{t-r(t)}^{t} \eta^T(s)Q_1 \eta(s)ds \leq \xi_1^T(t) \Psi_1 \xi_1(t) + h_2 \xi_1^T(t) M^T Q_1^{-1} M \xi_1(t),
\]

33
\[-\int_{t-\tau(t)}^{t-h_2} \eta^T(s)(Q_1 + Q_2)\eta(s)ds \leq \xi_1^T(t)\Psi_2\xi_1(t)(h_2 - h_1)\xi_1^T(t)N^TN + (Q_1 + Q_2)^{-1}N\xi_1(t),\]

\[-\int_{t-\tau(t)}^{t-h_1} \eta^T(s)Q_2\eta(s)ds \leq \xi_1^T(t)\Psi_3\xi_1(t) + (h_2 - h_1)\xi_1^T(t)S^TN_2^{-1}S\xi_1(t),\]

where

\[
\Psi_1 = \begin{bmatrix}
M_{11} + M_{11}^T & -M_{11} - M_{12}^T & M_{13}^T & M_{14}^T & M_{15}^T & M_{16}^T & M_{17}^T & M_{18}^T & M_{19}^T & M_{20}^T \\
* & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
0 & N_{11} & 0 & -N_{11} & 0 & 0 & 0 & 0 & N_1 & 0 \\
* & N_{12} + N_{12}^T & N_{13}^T & N_{12} + N_{14}^T & N_{15}^T & N_{16}^T & N_{17}^T & N_{18}^T & N_{19}^T & N_{20}^T
\end{bmatrix},
\]

\[
\Psi_2 = \begin{bmatrix}
0 & N_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & N_{13} & N_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & * & * & * & 0
\end{bmatrix},
\]

\[
\Psi_3 = \begin{bmatrix}
0 & S_{11} & S_{12} & S_{13} & S_{14} & 0 & 0 & 0 & 0 & 0 & 0 & S_1 \\
* & -S_{12} & S_{12} & S_{13} & S_{14} & 0 & 0 & 0 & 0 & 0 & 0 & S_2 \\
* & * & S_{13} + S_{13}^T & S_{14}^T & S_{15}^T & S_{16}^T & S_{17}^T & S_{18}^T & S_{19}^T & S_{20}^T \\
* & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_3 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & S_4 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 & S_5 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 & S_6 \\
* & * & * & * & * & * & * & 0 & 0 & 0 & S_7 \\
0 & S_{10} + S_{10}^T & S_{10} & S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} & S_{17} & S_{18} & S_{19} & S_{20}
\end{bmatrix},
\]

\[
M = \begin{bmatrix}
M_1^T & M_2^T & M_3^T & M_4^T & M_5^T & M_6^T & M_7^T & M_8^T & M_9^T & M_{10}^T \\
M_1 & M_2 & M_3 & M_4 & M_5 & M_6 & M_7 & M_8 & M_9 & M_{10}
\end{bmatrix},
\]

\[
N = \begin{bmatrix}
N_{11}^T & N_{12}^T & N_{13}^T & N_{14}^T & N_{15}^T & N_{16}^T & N_{17}^T & N_{18}^T & N_{19}^T & N_{20}^T \\
N_{11} & N_{12} & N_{13} & N_{14} & N_{15} & N_{16} & N_{17} & N_{18} & N_{19} & N_{20}
\end{bmatrix},
\]

\[
S = \begin{bmatrix}
S_{11}^T & S_{12}^T & S_{13}^T & S_{14}^T & S_{15}^T & S_{16}^T & S_{17}^T & S_{18}^T & S_{19}^T & S_{20}^T \\
S_{11} & S_{12} & S_{13} & S_{14} & S_{15} & S_{16} & S_{17} & S_{18} & S_{19} & S_{20}
\end{bmatrix},
\]

34
\[
\xi_1^T(t) = \begin{bmatrix}
    x^T(t) & x^T(t-\tau(t)) & x^T(t-h_1) & x^T(t-h_2) & \dot{x}^T(t) & f^T(x(t)) & f^T(x(t-\tau(t)))
\end{bmatrix}^T \\
\left(\int_{t-\tau(t)}^{t} x(s)ds\right)^T \left(\int_{t-h_2}^{t-\tau(t)} x(s)ds\right)^T \left(\int_{t-\tau(t)}^{t-h_1} x(s)ds\right)^T
\]

and \(\eta(s) = \begin{bmatrix} x^T(s) & \dot{x}^T(s)\end{bmatrix}^T\).

**Proof:** The proof of this Lemma immediately follows from Lemma 2 in [14].

**Lemma 1.16.15** For any matrices \(\tilde{R}_1 = \begin{bmatrix} R_1 & R_2 \\ R_2^T & R_3 \end{bmatrix} > 0\), \(\tilde{R}_2 = \begin{bmatrix} R_4 & R_5 \\ R_5^T & R_6 \end{bmatrix} > 0\), \(M_j, N_j, S_j, U_j, V_j, (j = 1, 2, \cdots, 6)\) and scalars \(\tau_0, \bar{\tau}\) and \(\alpha_0, 0 < \alpha_0 < 1\), the following inequalities hold

\[
- \int_{t-\alpha_0\tau_1(t)}^{t} \eta^T(s) \tilde{R}_1 \eta(s) ds \leq \xi_2^T(t) \tilde{\Psi}_1 \xi_2(t) + \alpha_0 \tau_0 \xi_2^T(t) M^T \tilde{R}_1^{-1} M \xi_2(t), \\
- \int_{t-\tau_1(t)}^{t} \eta^T(s) \tilde{R}_1 \eta(s) ds \leq \xi_2^T(t) \tilde{\Psi}_2 \xi_2(t) + \tau_0 (1 - \alpha_0) \xi_2^T(t) N^T \tilde{R}_1^{-1} N \xi_2(t), \\
- \int_{t-\tau_0}^{t} \eta^T(s) \tilde{R}_1 \eta(s) ds \leq \xi_2^T(t) \tilde{\Psi}_3 \xi_2(t) + \tau_0 \xi_2^T(t) S^T \tilde{R}_1^{-1} S \xi_2(t), \\
- \int_{t-\tau_2(t)}^{t} \eta^T(s) \tilde{R}_2 \eta(s) ds \leq \xi_2^T(t) \tilde{\Psi}_4 \xi_2(t) + (\bar{\tau} - \tau_0) \xi_2^T(t) U^T \tilde{R}_2^{-1} U \xi_2(t), \\
- \int_{t-\bar{\tau}}^{t} \eta^T(s) \tilde{R}_2 \eta(s) ds \leq \xi_2^T(t) \tilde{\Psi}_5 \xi_2(t) + (\bar{\tau} - \tau_0) \xi_2^T(t) V^T \tilde{R}_2^{-1} V \xi_2(t),
\]

where

\[
\tilde{\Psi}_1 = \begin{bmatrix}
    M_4 + M_7^T & -M_5 + M_7^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    -M_6 + M_7^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & * \\
    * & * & * & * & * & * & * & * & * & *
\end{bmatrix}
\]
\[ \bar{\psi}_2 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & N_4 + N_T^T & -N_4 + N_T^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & -N_5 - N_5^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & -N_4 + N_T^T & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & * & * & * & 0 \\
\end{bmatrix} \\
\]

\[ \bar{\psi}_3 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & S_4 + S_T^T & -S_4 + S_T^T & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & -S_5 - S_5^T & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & * & * & * & 0 \\
\end{bmatrix} \\
\]

\[ \bar{\psi}_4 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & U_4 + U_T^T & -U_4 + U_T^T & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & -U_5 - U_5^T & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & * & * & * & 0 \\
\end{bmatrix} \\
\]

\[ \bar{\psi}_5 = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & V_4 + V_T^T & V_4^T - V_4 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & -V_5^T - V_5 & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & 0 & 0 & 0 & 0 \\
* & * & * & * & * & * & * & 0 & 0 & 0 \\
* & * & * & * & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & * & * & * & 0 \\
\end{bmatrix} \\
\]

\[
M = \begin{bmatrix}
M_T^T & M_T^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
M_4^T & M_4^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
M_5^T & M_5^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
M_3^T & M_3^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
M_6^T & M_6^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \\
\]

\[
N = \begin{bmatrix}
0 & N_T^T & N_T^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_4^T & N_5^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_5^T & N_6^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_5^T & N_6^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_4^T & N_5^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & N_5^T & N_6^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \\
\]

36
\[
S = \begin{bmatrix}
0 & 0 & S_1^T & S_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_3^T & 0 & 0 \\
0 & 0 & S_2^T & S_2^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & S_4^T & 0 & 0 \\
\end{bmatrix},
\]
\[
U = \begin{bmatrix}
0 & 0 & 0 & U_1^T & U_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & U_3^T & 0 \\
0 & 0 & 0 & U_4^T & U_4^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & U_6^T & 0 \\
\end{bmatrix},
\]
\[
V = \begin{bmatrix}
0 & 0 & 0 & 0 & V_1^T & V_1^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & V_3^T \\
0 & 0 & 0 & 0 & V_4^T & V_4^T & 0 & 0 & 0 & 0 & 0 & 0 & 0 & V_6^T \\
\end{bmatrix},
\]
\[
\xi(t) = x(t) x^T(t_0) x(t - \tau_1(t)) x(t - \tau_1(t)) x(t - \tau(t)) x(t - \tau(t)) x^T(t - \tau(t)) x^T(t)
\]
\[
f^T(x(t)) f^T(x(t - \tau_1(t))) f^T(x(t - \tau_2(t))) \int_{t-\tau_1(t)}^{t} x^T(s) ds \int_{t-\tau(t)}^{t-\tau_1(t)} x^T(s) ds \int_{t-\tau_2(t)}^{t-\tau(t)} x^T(s) ds \int_{t-\tau(t)}^{t} x^T(s) ds
\]
\[
\int_{t-\tau_0}^{t-\tau(t)} x^T(s) ds \int_{t-\tau_0}^{t-\tau(t)} x^T(s) ds \int_{t-\tau(t)}^{t} x^T(s) ds \int_{t-\tau(t)}^{t} x^T(s) ds
\]

**Proof:** The proof of this Lemma immediately follows from Lemma 2 in [14].

**Definition 1.16.16 (Itô’s formula)** Let \( x(t) \) be a \( d \)-dimensional Itô process on \( t \geq 0 \) with the stochastic differential

\[
dx(t) = f(t)dt + g(t)dB(t),
\]

where \( f \in L^1(\mathbb{R}; \mathbb{R}^d) \) and \( g \in L^2(\mathbb{R}; \mathbb{R}^{d \times m}) \). Let \( V \in C^{2,1}(\mathbb{R}^d \times \mathbb{R}; \mathbb{R}) \). Then \( V(x(t), t) \) is again an Itô process with the stochastic differential given by

\[
dV(x(t), t) = LV(x(t), t)dt + V_x(x(t), t)g(t)dB(t)
\]

where

\[
LV(x(t), t) = V_t(x(t), t) + V_x(x(t), t)f(t) + \frac{1}{2} \text{trace} \left( g^T(t)V_{xx}(x(t), t)g(t) \right)
\]

\[
V_t(x(t), t) = \frac{\partial V(x(t), t)}{\partial t}, V_x(x(t), t) = \left( \frac{\partial V(x(t), t)}{\partial x_1}, \frac{\partial V(x(t), t)}{\partial x_2}, ..., \frac{\partial V(x(t), t)}{\partial x_d} \right),
\]

and

\[
V_{xx}(x(t), t) = \left( \frac{\partial^2 V(x(t), t)}{\partial x_i \partial x_j} \right)_{d \times d}
\]

37
1.17 Thesis outline and contributions overview

This thesis is dedicated to investigate the stability analysis of NNs with different time-delays. Different kinds of NNs like HNNs, SNNs and neutral type NNs are considered. The time-varying delays are assumed to be bounded by a positive scalar. The main objectives of this thesis are to propose less conservative conditions to the global stability problems and to develop efficient algorithms to deal with the global stability problems of DNNs. To the best of our knowledge, this is the first attempt to discuss the global stability of DNNs in terms of LMIs. The feasible matrices can be efficiently accomplished by resorting standard numerical softwares and the optimal scalar values can be obtained by solving corresponding convex optimization problems.

Chapter 2 deals the stability analysis for SNNs with time-delays. Delay dependent stability conditions are developed for SNNs. The discrete delay is assumed to be time-varying and belonging to a given interval, which means that the lower and upper bounds of interval time-varying delays are available. The designed criterion is formulated in terms of LMI and can be easily checked via the Matlab LMI control toolbox.

In Chapter 3, the delay-dependent stability analysis is considered for neutral type SNNs with polytopic and linear fractional uncertainties is investigated. The delay-interval dependent robust stability criteria is studied for neutral SNNs under polytopic and linear fractional uncertainties. Moreover, the lower and upper bounds of the delay interval are assumed to be known. Based on a suitable Lyapunov-Krasovskii functional and Ito’s differential rule, some delay-dependent stability conditions are derived for the considered neutral SNNs in terms of LMIs. Numerical examples are provided to illustrate the theory.

In Chapter 4, the delay-dependent stability problem is studied for a class of NNs with Markovian jumping parameters and multiple delays that include discrete and neutral delays. By constructing a new Lyapunov-Krasovskii functional and employing some analysis techniques, sufficient conditions are derived for the considered system.
in terms of LMIs, which can be easily calculated by Matlab LMI control toolbox. Numerical examples are given to illustrate the effectiveness of the proposed method.

Chapter 5 is focused on the delay-probability-distribution-dependent robust stability problem for a class of uncertain SNNs with time-varying delays. The information of probability distribution of the time-delay is considered and transformed into parameter matrices of the transferred SNNs model. Based on the Lyapunov-Krasovskii functional and stochastic analysis approach, a novel delay-probability-distribution-dependent sufficient condition is obtained in the LMI form such that delayed SNNs are robustly globally asymptotically stable in the mean square for all admissible uncertainties. An important feature of the result is that the stability conditions are dependent on the probability distribution of delays and upper bound of the derivative is allowed to be greater than or equal to 1. Numerical examples are given for the comparison to illustrate the effectiveness of the derived results.