Doctoral Dissertation

Continuous Time Structural Autoregressive and Autoregressive Moving Average Modeling Approaches and Their Application

Marina Demeshko

March 2015

Graduate School of Engineering

Osaka University

Preface

This dissertation presents my research on techniques of canonical Autoregressive and Autoregressive Moving Average modeling from multivariate data. The dissertation is the result of the research during the Ph.D. course at the Department of Information and Communication Technology, Graduate School of Engineering, Osaka University. The dissertation is organized as follows.

Chapter 1 describes the background, the motivation, the purpose of this research, and the outline of this dissertation. The key objective of this dissertation is to construct methodologies for finding the structure of the dependencies among the multiple processes in the objective system.

Such techniques allow us to infer the unknown data generating mechanism or to model the target data with an efficient manner. For the purpose, we focus on the Autoregressive and then on more complex Autoregressive Moving Average models. These two models form the basis of the dissertation, which we further extend in the upcoming chapters.

Chapter 2 is devoted for the Continuous time Structural Vector Autoregressive (CSVAR) modeling approach. We also provide its theoretical and numerical justifications.

In Chapter 3, we describe a Continuous time Structural Autoregressive Moving Average (CSARMA) modeling approach which is essential for the analysis of more complex systems. The validity of the proposed method is verified through numerical simulations and also on an application to the real-world system.

Chapter 4 describes an application of the developed CSARMA method to the nuclear reactor stability analysis through the investigation of the anomaly unstable event during the operation of Kernkraftwerk Boiling Nuclear Reactor.

Chapter 5 concludes this dissertation.

Acknowledgment

I worked on this dissertation under the supervision of Prof. Takashi Washio. I would first like to thank him for sharing his knowledge, pushing me to work hard and constantly try to improve my work.

I would also like to thank my other co-authors for their help in letting me more productive that I would have been able to on my own: Dr. Yoshinobu Kawahara and Dr. Shohei Shimizu in Osaka University, Dr. Hakim Ferroukhi and Dr. Abdelhamid Dockhane in Paul Scherrer Institut, and Dr. Yuriy N. Pepyolyshev in Joint Institute of Nuclear Research.

I'd like to express my deep gratitude to to Joint Institute of Nuclear Research in Dubna and to Kernkraftwerk in Leibstadt for sharing their data on nuclear reactors.

I would like to acknowledge Prof. Kazunori Komatani, Prof. Noboru Babaguchi, Prof. Kyou Inoue, Prof. Kenichi Kitayama, Prof. Seiichi Sampei and Prof. Tetsuya Takine from Graduate School of Engineering at Osaka University.

I also had several supports from all of laboratory members. In particular, I would like to thank Ms. Hiroko Okada for her clerical assistance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Many real world systems are well approximated by continuous time multivariate linear Markov processes. They are mostly observed in form of multivariate time series data sampled at discrete time steps for their digital processing, and we use discrete time parametric model approximations to analyze the systems through the observed data. Discrete time Multivariate or Vector AR (DVAR)[1] and Autoregressive and Moving Average (DARMA) models are used for such analyses when the objective system is stable[2], controllable[3] and observable[4] (Brockwell & Davis, 1991). DVAR modeling is more popular, because the parameter estimation procedure for it is much easier than for the multivariate ARMA model (Kizilkaya & Kayran, 2006). However, when an objective continuous time linear Markov system is not exactly described by an AR process, the DVAR model with finite order is just an approximation of the discrete time multivariate linear Markov system in many cases, while the DARMA model is a general exact notion of it. Therefore, if our objective system behind a given data set is well approximated by a multi-

[1] In the past studies, the notion of a vector autoregressive model (VAR) is usually used as a discrete time model. In our study, to separate the vector autoregressive processes for discrete and continuous time domains, we will use the term of a discrete time vector autoregressive (DVAR) model instead.

[2] The system is stable if all nearby initial conditions converge to the equilibrium point.

[3] The system is controllable if its current state can be transfered to any given state by applying an appropriate input series over a finite time period.

[4] The system is observable if, for any possible sequence of state and control vectors, the current state can be determined in finite time using only the outputs.

variate linear Markov process, we usually should use DARMA modeling to identify the structure and the parameters of the system. Accordingly, the DARMA modeling has been widely used in many fields, *e.g.*, biology (Perrott & Cohen, 1996), medicine (Gannabathula, 1988), finance (Shittu & Yaya, 2009), economy (Chen et al., 2010), physics (Kuroda et al., 1985; Tran, 1992, 2003), engineering (Zhao et al., 2007) and energy (Zhang & Xu, 2010).

However, the DVAR/DARMA models have some drawback to identify the system dynamics, *i.e.*, the lack of canonicality of their model representation. A canonical model is a unique system representation which is physically or mathematically admissible. Both the DVAR and the DARMA models are not canonical, because an objective system can be represented by infinitely many DVAR/DARMA models as explained in the later chapters. This comes from the fact that they cannot represent the system dynamics in time scales shorter than their discretized time interval. This limitation induces difficulty in identifying the structure and the parameters of the objective system dynamics. In this study, our goal is to accurately derive a canonical representation of an objective system behind a given time series dataset under a generic assumption which does not virtually limit its applicability when the system is well approximated by a continuous time, multivariate and linear Markov process.

Many studies have been conducted to derive Structural AR/ARMA (SVAR/SARMA) models in discrete time domain (Hyvarinen et al., 2008; Gottschalk, 2001; Pfaff & Kronberg, 2008; Kilian, 2011; Kawahara et al., 2011; Mainassara & Francq, 2009). The term structural denotes that the model represents the directions of dependencies between variables, *i.e.*, which variable depends on which variable, in addition to the quantitative dynamics of an objective system. The structurality and the canonicality of the models are closely related. Under the axiomatic believe that the directions of the dependencies between variables in a physical system should be unique, only a unique SVAR/SARMA model is admissible. Therefore, the SVAR/SARMA models are canonical.

However, the SVAR/SARMA models have not been empirically derived from observed time series without introducing some strong assumptions by using the standard DVAR/DARMA modeling methods, because the DVAR/DARMA models lack

their canonicality. Therefore, some extra information related to the depenency between the variables have to be introduced to the modeling. Most past studies have been concentrated on the structural vector autoregressive (SVAR) models, which are a subclass of the SARMA models because of the aforementioned advantage of the AR modeling on the parameter estimation. In these studies, identification of the SVAR model requires some strong assumptions on their system structure and/or variables. For example, a study introduced assumptions of acyclic dependency among variables and non-Gaussian external noises (Hyvarinen et al., 2008). Other methods also require some specific assumptions, such as orthogonality and non-linear restrictions on parameter matrices of the external noises, and recursive ordering of system parameters (Gottschalk, 2001; Pfaff & Kronberg, 2008; Kilian, 2011). We should note that most of these approaches have not been extended to more generic multivariate ARMA modeling. Only few studies investigated SARMA modeling, but they also require some strong assumptions for model identification, such as the acyclicity of an objective system and non-Gaussianity of its external noises (Kawahara et al., 2011) and system representation by a small number of parameters and partial independence of the noises (Mainassara & Francq, 2009). However, the introduction of these assumptions focusing on some special systems largely limits the applicability of SVAR and SARMA modeling.

On the other hand, the CVAR/CARMA model, which consists of differential equations in continuous time domain, is known to be canonical. In contrast to aforementioned DVAR/DARMA model, the CVAR/CARMA model is an exact representation of the continuous AR/ARMA processes of the objective system, *i.e.*, it reflects all system dynamics. This characteristic does not admit its multiple representations of the system as explained in the latter chapters. Therefore, the CVAR/CARMA model is a uniquely admissible representation, and thus it is canonical.

In this dissertation, we aim to develop a technique that do not require any strong assumptions to derive the canonical models, *i.e.*, the SVAR/SARMA and the CVAR/CARMA models. We first derive new mathematical constraints on the SVAR/SARMA model from the canonicality of the CVAR/CARMA model under the assumption that the objective system is well approximated by a con-

tinuous multivariate and linear Markov process. Subsequently, we propose novel methods to derive the SVAR/SARMA and the CVAR/CARMA models from the estimated DVAR/DARMA models based on the mathematical constraints. The use of the canonicality of CVAR/CARMA model in continuous time domain for the SVAR/SARMA modeling in discrete time domain has not been studied yet, and this is our novel idea originated in this dissertation work. We name these new methods as Continuous and Structural Vector Autoregressive modeling (CSVAR) for the VAR model and Continuous and Structural Autoregressive Moving Average modeling (CSARMA) for the ARMA model, respectively.

We further apply these techniques to a real world problem. Both the CSVAR and CSARMA methods are expected to be strong instruments for the analysis of the dynamics dependency between variables in complex engineering systems. In this dissertation, we concentrate on an application to Boiling Water Reactors (BWRs). Since the early stage of the technology development for BWRs, it has been recognized that unstable power oscillations could occur due to neutron and thermal-hydraulic coupling in the core dynamics. Hence, the reactors have been designed with a large stability margin and operated conservatively so as not to encounter instability. However, instability events can still occur at the BWR operation. One of such events was observed at the stability test during start-up of Kernkraftwek Leibstadt (KKL) Cycle 24 (Ferroukhi, 2008). The COSMOS on-line stability monitoring system (Blomstrand et al., 1993) indicated the steam flow's fluctuations with unexpectedly high level amplitudes. Consequently, the reactor was stopped by the emergency system. That lessened the efficiency of the reactor operation. The event was unexpected, since the core was operated at low power and low steam flow conditions.

To prevent reappearance of such an event for the reactor's safety management, it is important to know the operation conditions that may affect its stability. Stability analysis methods traditionally used are categorized into non-parametric and parametric methods. The non-parametric methods mainly provide information on reactor stability based on the analysis of power spectral densities and spectral coherence of the reactor signal fluctuations. The parametric analysis methods are used to evaluate the stability of the reactor operation by estimating stability mea-

sures such as decay ratio (DR) and natural frequency (FR) of the reactor power, or the Jacobian of the objective system. However, all these methods do not give us clear knowledge on interactions among individual reactor processes causing the reactor instability.

For understanding the mechanism of the reactor during an instability event, we need to have a mathematical model and its corresponding parameters that precisely reflect the reactor processes. Since the reactor noise generation processes can be often approximated by a linear Markov system and they are observed as a time series sampled at discrete time steps, we often apply Discrete time Autoregressive Moving Average (DARMA) modeling to obtain the model and the parameters in the framework of the parametric analysis. However, since a multivariate DARMA model is not canonical, *i.e.*, it does not uniquely represent the objective reactor processes, it makes difficult to understand the dynamics of the objective system and its individual processes.

In this dissertation, we expect that an application of a novel CSARMA approach (Demeshko et al., 2013, 2014) to the analysis of the instability event at the nuclear plant Kernkraft Leibstadt (KKL) will allow us to investigate the dependency among the processes in the system and give us more insights on the physical mechanisms of the system.

Based on these considerations, we aim to achieve the following main research objectives of the dissertation.

(1) Upon generic assumptions on the objective continuous time, multivariate and linear Markov system, we theoretically discover the constraints on the SVAR model by applying the canonicality of the CVAR model. By using these mathematical constraints, we establish a new modeling approach to derive the CVAR model and the SVAR model from the DVAR model estimated by using time series data observed from the system.

(2) We further extend the discovered mathematical constraints on the SVAR model to the SARMA model derived from the canonicality of the CARMA model. By using the new mathematical constraints, we propose a new modeling approach to derive the multivariate CARMA model and the SARMA

model from the multivariate DARMA model estimated by using time series data observed from the system.

(3) We apply the developed modeling methods to the data produced from the real world problem, *i.e.*, the instability event at the Kernkraftwerk Leibstadt nuclear plant. We obtain new insights on the dependency between variables of nuclear reactor's physical processes by applying the CSARMA approach.

In the rest of this dissertation, all of the CARMA, the SARMA and the ARMA models are multivariate unless we explicitly state their univariateness.

1.2 Summary of contributions

We briefly summarize the contributions of each chapter:

- **Chapter 2**: The main theoretical contribution of this chapter is the derivation of the mathematical constraints from the canonicality of the CVAR model to make a DVAR model structural. We could derive such constraints through the time discretization of the CVAR model.

  The second contribution of our study is the establishment of a novel modeling approach called Continuous time Structural Vector Autoregressive (CSVAR) modeling that is based on the derived constraints. It allows us to derive the CVAR and the SVAR models from a given DVAR model.

- **Chapter 3**: The main theoretical contribution of this chapter is the extension of the mathematical constraints on the SVAR model to those of the SARMA model by applying the time discretization of the CARMA model.

  Based on these constraints, we achieved our second contribution, *i.e.*, we extend the CSVAR modeling to the modeling approach that allows us to derive the CARMA and the SARMA models from a given DARMA model. This technique is named a Continuous time Structural Autoregressive Moving-Average (CSARMA) modeling approach.

- **Chapter 4**: We apply the developed CSARMA modeling approach for the nuclear reactor stability analysis. We investigate anomalous instability of Boiling Water Reactor (BWR) of Kernkraftwerk nuclear station, that was observed during event Cycle 24. The main result of our analysis provides the clear and valid dependencies between variables of the system, and give us more insights on the system dynamics in comparison with Signal Transmission Path analysis which is a representative conventional analysis technique.

1.3 Relationship of three main chapters

We explain the relations between the ideas introduced in three main chapters of the dissertation.

Chapter 2: CSVAR modeling approach

In this chapter, we derive mathematical constraints on the SVAR model by using the discrete time approximation of the CVAR model. By applying these constraints, we clarified the mathematical relations between the CVAR, the SVAR and the DVAR models of the system and developed the CSVAR modeling approach. Chapter 2 provides the mathematical constraints and the modeling approach to be extended in Chapter 3.

Chapter 3: CSARMA modeling approach

Based on the idea presented in Chapter 2, we extended the mathematical constraints discovered in Chapter 2 to include the moving-average part of the ARMA model. The constraints derive new mathematical relations between the CARMA, the SARMA and the DARMA models. These relations enable us to establish the CSARMA modeling approach. This approach developed in Chapter 3 is applied to demonstrate its practicality in a real world problem in Chapter 4.

Chapter 4: Analysis of BWR instability mechanisms applying a CSARMA approach

In this chapter, we present practical applicability of the CSARMA method, developed in Chapter 3, to a real world problem, *i.e.,* an instability analysis of a nuclear power plant. Furthermore, we provide new insights, derived by the CSARMA method, on the instability mechanism of the plant.

Chapter 2

CSVAR modeling approach

2.1 Introduction

Usually real world systems are observed in a way of time series data sampled at discrete time steps for the digital processing. In case when such an objective system is well represented by continuous time, multivariate, linear Markov, stable and controllable system, a discrete time vector autoregressive (DVAR) model is often used to analyze it (Brockwell & Davis, 1991). However, according to previous studies (Stamer et al., 1996; Gottschalk, 2001), the DVAR model is not canonical, *i.e.*, it does not have a bijective correspondence with the objective system for which it is derived. This induces difficulty in identifying the structure and the parameters of the objective system dynamics.

As a remedy to this limitation of the DVAR modeling, a structural vector AR (SVAR) model (Moneta, Entner, Hoyer, & Coad, 2010) has been studied. The SVAR model has a bijective correspondence with a unique system. It is used for the modeling when an objective continuous time linear Markov system is exactly described by an AR process. Moreover the SVAR model provides information on the propagation of influences among variables in the system, even if the DVAR is only an approximation of the objective dynamic system. However, in the past studies, some strong assumptions were required on the system structure and/or variables for the identification of the SVAR model. These assumptions were acyclic dependency among variables and non-Gaussianity of external noises (Hyvarinen et al., 2008), orthogonality and non-linear restrictions on parameter matrices of the external noises (Gottschalk, 2001; Pfaff & Kronberg, 2008; Kilian, 2011), and recursive ordering of system parameters (Gottschalk, 2001; Kilian, 2011). They largely limit

the applicability of the SVAR modeling in the system structure analysis.

Based on these considerations, we aim to achieve the following four research objectives in this chapter.

(2.1) Show under a generic assumption, that the DVAR model of a continuous time, multivariate, linear Markov system is canonical.

(2.2) Clarify mathematical relations among a continuous time vector AR (CVAR) model, a SVAR model and a DVAR model of the system.

(2.3) Present a new approach named Continuous time Structural Vector Autoregressive (CSVAR) modeling for the CVAR and the SVAR models based on the DVAR model obtained from observed time series data.

(2.4) Demonstrate using numerical experiments on artificial and real world time series the applicability and the accuracy of the proposed approach.

Based on the above arguments this chapter is organized as follows. Section 2.2 explains past studies related with our work and clarifies technical issues to be addressed in this chapter. In Section 2.3, we established a new principle to reconstruct the CVAR and SVAR models from the DVAR model estimated from observed time series. In Section 2.4, we show a performance of our proposed approach by numerical experiments using artificial and real world data and its comparison with a past representative SVAR modeling approach. Section 2.6 concludes this part.

2.2 Preliminary discussion and analysis

2.2.1 Related work

A DVAR model Eq.(2.1) of order p is a general representation of a stable, controllable, multivariate, liner Markov system with a given d-dimensional variable vector $Y(t)$ observed with a discrete time step Δt.

$$Y(t) = \sum_{j=1}^{p} \Phi_j Y(t - j\Delta t) + U(t), \tag{2.1}$$

where Φ_j are $d \times d$ coefficient matrices. $U(t)$ is a d-dimensional unobserved noise vector that is i.i.d. in a discrete time domain (Brockwell & Davis, 1991). This model has ambiguity to represent infinitely many systems, since Eq.(2.1) is equivalent to

$$Y(t) = \sum_{j=1}^{p} \Phi_j Y(t - j\Delta t) + QW(t), \tag{2.2}$$

where Q is a $d \times d$ regular matrix and $W(t) = Q^{-1}U(t)$. Thus, depending on the choices of Q, $Y(t)$ has infinitely many impulse responses for $W(t)$ (Moneta et al., 2010). In other words, for a given observed time series, $Y(t)$, the infinite number of systems that could generate $Y(t)$ are represented by a unique DVAR model with parameters Φ_j, $j = 1, \ldots, p$. Therefore, the DVAR model has no bijective correspondence to the objective system, *i.e.*, it is not canonical. Thus, we need to know a unique combination of Q and $W(t)$ which corresponds to the objective system dynamics in a bijective manner to identify the system uniquely. Having such Q and $W(t)$, we multiply Q^{-1} from the left-hand-side of Eq.(2.2) as

$$Q^{-1}Y(t) = \sum_{j=1}^{p} Q^{-1}\Phi_j Y(t - j\Delta t) + W(t).$$

By adding $Y(t)$ to both hand sides and moving $Q^{-1}Y(t)$ to the right-hand-side, we obtain the following formula.

$$Y(t) = (I - Q^{-1})Y(t) + \sum_{j=1}^{p} Q^{-1}\Phi_j Y(t - j\Delta t) + W(t).$$

This is further rewritten as

$$Y(t) = \sum_{j=0}^{p} \Psi_j Y(t - j\Delta t) + W(t), \tag{2.3}$$

where $\Psi_0 = I - Q^{-1}$, $\Psi_j = Q^{-1}\Phi_j$. This equation includes fast effects in its representation by a matrix Ψ_0 representing a feedback of $Y(t)$ on itself. If there are no such effects in the objective system, then Ψ_0 should be a zero matrix, otherwise Ψ_0 is non-zero and $Q \neq I$. By deriving a unique matrix Ψ_0, we define the unique combination of Q and $W(t)$ in Eq.(2.3), which provides a unique model with parameters Ψ_j, $j = 0, \ldots, p$ that bijectively represents the unique objective system. Therefore,

this model is canonical. In the past studies, Eq.(2.3) is called a Structural Vector Autoregressive (SVAR) model (Moneta et al., 2010; Kawahara et al., 2011).

The SVAR model is provided by the derivation of the unique Ψ_0. As shown in Eq. (2.3), matrices Ψ_0 and Ψ_j directly depend on Q. However, matrix Q is not reproduced from the DVAR model, because it is integrated in Φ_j as $\Phi_j = Q\Psi_j$. Accordingly, we need some extra information to provide Q. However, even if we introduce an orthonormality constraint $QQ^T = E_t(U(t)U(t)^T)$ to make noises in $W(t)$ mutually uncorrelated, the representation of the DVAR model given by this approach is not unique, because there are many choices for Q which satisfy that constraint, *e.g.*, QO satisfying $QO(QO)^T = QOO^TQ^T = QQ^T$, where O is any orthonormal matrix (Moneta et al., 2011).

There are many studies on this issue for identifying the SVAR model. Most of these studies require strong assumptions on the objective system. For example, a study (Gottschalk, 2001) introduced a constraint named *exclusion restriction* together with the orthonormality of the external noises. It requires some domain knowledge of the objective system to obtain the order of the noises in $U(t)$ and hence their corresponding variables in $Y(t)$. The order of the noises is further used for Cholesky decomposition that uniquely derives a strictly lower triangular Q. A study (Kilian, 2011) introduced another constraint named a sign restriction. It also requires some domain knowledge on the signs of some elements of Q with the aforementioned orthonormality of the noise vector to select the matrix Q. Such Q is searched by testing the sign restriction after randomly generating orthonormal matrices with QR decompositions. A study (Hyvarinen et al., 2008) introduced less domain specific assumptions on the system. It proposed a novel method named AR-LiNGAM by assuming acyclic dependency among variables in $Y(t)$ and non-Gaussianity of external noises in $W(t)$ in addition to mutual independence and temporal uncorrelation of $W(t)$. The ordering information of the noises in $W(t)$ and their corresponding variables in $Y(t)$ were further provided by the application of independent component analysis (ICA) to least-square residual noises of the DVAR modeling.

However, in many real-world applications, the aforementioned assumptions on the objective system significantly limit the applicability of these approaches, since

such firm domain knowledge is not readily available. In this regard, we require a more generic approach to the SVAR modeling without using strong assumptions. To address this issue, we propose a novel modeling approach that uses only a very general assumption which subsumes stability and controllability of the objective system.

2.2.2 Canonicality of VAR models

In (Pearl, 2000), Pearl stated that each noise in a canonical model has its own unique variable to directly and instantly change within a negligibly short time period required by attaining the equilibrium. Similar characterization and analysis of the canonical models can be seen in the literatures (Fisher, 1970; Iwasaki & Simon, 1994; Lacerda et al., 2008; Mooij et al., 2013). $W(t)$ in the SVAR model Eq.(2.3) holds the characteristics of the noise vector in the canonical model, because each noise in $W(t)$ has its unique variable in $Y(t)$ to directly and instantly change. The term of Ψ_0 represents fast effects in $Y(t)$ which occur within the sampling time interval Δt. The other AR terms represent components generated by the system dynamics within a finite time interval, if the system is stable. Accordingly, the SVAR model is canonical under a given combination of regular matrix Q and $W(t)$ that correspond to the objective system.

Comparing the DVAR and the SVAR models, we obtain the following relations among their matrices and vectors (Kawahara et al., 2011; Hyvarinen et al., 2008).

$$\Phi_j = (I - \Psi_0)^{-1}\Psi_j, \tag{2.4a}$$

$$\Psi_j = (I - \Psi_0)\Phi_j, \tag{2.4b}$$

$$U(t) = (I - \Psi_0)^{-1}W(t), \tag{2.5a}$$

$$W(t) = (I - \Psi_0)U(t). \tag{2.5b}$$

We see that given a combination of $\Phi_j (j = 1, \ldots, p)$, there are various combinations of $\Psi_j (j = 0, \ldots, p)$, which induce the combination of $\Phi_j (j = 1, \ldots, p)$ through Eq.(2.4a). Thus, there exist multiple SVAR models, which induce the DVAR model. On the other hand, given a combination of Ψ_0, Ψ_j $(j = 1, \ldots, p)$, Q is uniquely

provided by $Q = (I - \Psi_0)^{-1}$, and a unique combination of $\Phi_j (j = 1, \ldots, p)$ is also provided through Eq.(2.4a). Thus, a given SVAR model induces only a unique combination of the DVAR model and the regular matrix Q. In other words, the correspondence from the SVAR model to the DVAR model is surjective. This is because the DVAR model is not canonical.

On the other hand, when a stable, controllable, multivariate, linear Markov system in a continuous time domain is approximated by an AR process, it is represented by a Continuous Time Vector AR (CVAR) model in Eq.(2.6) consisting of continuous time stochastic differential equations with noise terms (Stamer et al., 1996).

$$Y^{(p)}(t) = \sum_{m=0}^{p-1} S_m Y^{(m)}(t) + W(t), \tag{2.6}$$

where $Y(t)$ and $W(t)$ are a d-dimensional observed variable vector and a d-dimensional external noise vector that is i.i.d. in a continuous time domain, respectively. $Y^{(m)}(t)$ is the m-th time differential of $Y(t)(= Y^{(0)}(t))$, and $S_m (m = 0, \ldots, p-1)$ is a $d \times d$ AR matrix (Stamer et al., 1996). Since this is a continuous time model, it includes the fast effects in $Y(t)$ which occur within a period smaller than Δt. In this model, each noise in $W(t)$ directly and instantly changes the highest time differential of its corresponding unique variable in $Y(t)$, and the change propagates to the other lower order time differentials. In this regard, $W(t)$ satisfies the character of the noise vector in a canonical model. Additionally, $W(t)$ in Eq.(2.6) is a unique noise vector, because application of $E(t) = P^{-1}W(t)$ with a $d \times d$ regular matrix P transforms Eq.(2.6) as

$$Y^{(p)}(t) = (I - P^{-1})Y^{(p)}(t) + \sum_{m=0}^{p-1} P^{-1} S_m Y^{(m)}(t) + E(t).$$

Because any instantaneous process to change the highest order differentials $Y^{(p)}$ by itself is not admissible in a complete process dynamics, $P \neq I$ is not admitted. Thus, $W(t)$ is unique. This is because of the fact that the CVAR model is an exact representation of the continuous AR processes in the objective system and thus includes all process dynamics. Accordingly, the CVAR model Eq.(2.6) is canonical, and the SVAR model Eq.(2.3) and the CVAR model Eq.(2.6) bijectively correspond

to the system dynamics. Since they represent the same stable, multivariate linear Markov system, which is controllable in both continuous and discrete time domains, the SVAR and the CVAR models have a bijective correspondence.

2.3 Proposed principle

In this section, we concentrate on objectives (2.1), (2.2) and (2.3) stated in Section 2.1. Firstly, we introduce an assumption that is necessary to derive mathematical relations between CVAR, SVAR and DVAR models and to provide theoretical bases for SVAR and CVAR modeling from given time series data. Then we show a bijective correspondence between all three models. Finally, we propose a new canonical modeling approach of the objective continuous time, multivariate, linear Markov system.

The proposed modeling principles and algorithm require the following assumption.

Assumption 1. *Given the CVAR model in Eq.(2.6) representing the objective system and a positive real constant $\Delta t > 0$ which is a sampling interval of time series data for the modeling, $\sum_{m=0}^{p} S_m \Delta t^{p-m}$ is a regular matrix, where $S_p = -I$ according to Eq.(2.6).* ■

As shown in the following lemma, given an objective, continuous time, multivariate, linear Markov system, this assumption always holds, if the system is stable and controllable.

Lemma 1. *Given the CVAR model in Eq.(2.6) representing the objective system with a positive real constant $\Delta t > 0$, $\sum_{m=0}^{p} S_m \Delta t^{p-m}$ is a regular matrix where $S_p = -I$, if the system is controllable and stable.* ■

Proof. *The proof is presented in Appendix 2.7.1.* □

Assumption 1 does not apply any essential limitation to the canonical modeling of the system as long as we use the DVAR modeling, because the stability and the controllability of the objective system are required for the estimation of its valid DVAR model by using time series data observed from the system (Brockwell & Davis, 1991).

To obtain the relations among the three models, first we concentrate on the canonical SVAR and CVAR models. As we discussed in previous section, they have a bijective correspondence. However, the variables of the CVAR model are in continuous time domain, while those of the SVAR model are in discrete time domain. Accordingly, we find their mathematical relations by applying a time discretization approximation to the CVAR model. One of the most representative approximation schemes for the time discretization is the backward higher order finite difference, which is a natural extension of Euler formula for the m-th derivative (Levy & Lessman, 1992). We introduce the following assumption to apply the time discretization approximation.

Assumption 2. *The approximation error by introducing the following backward higher order finite difference is sufficiently small for the time discretization of the CVAR model.*

$$Y^{(m)}(t) \simeq \frac{1}{\Delta t^m} \sum_{j=0}^{m} (-1)^j \frac{m!}{(m-j)!j!} Y(t - j\Delta t). \tag{2.7}$$

■

This finite difference scheme is consistent, *e.g.*, the finite difference equations derived by Eq. (2.7) converge to their original differential equation when $\Delta t \to 0$. The approximation error is $O(Y^{(m+1)}(t)\Delta t)$, which converges to zero for $\Delta t \to 0$. Moreover, the convergence of the finite difference equations' solutions provided by Eq. (2.7) to that of the original time differential equation is ensured by Lax-Richtmyer theorem, since the objective system is linear and stable (Lax & Richtmyer, 1956; Strikwerda, 1989).

Lemma 2. *The finite difference approximation of the CVAR model using Eq.(2.7) is unbiased under Assumption 1.* ■

Proof. *The proof is presented in Appendix 2.7.2.* □

Therefore, we obtain the following lemma.

Lemma 3. *Under Assumptions 1 and 2, a discrete time approximation of the CVAR model is represented as follows.*

$$\begin{aligned} Y(t) = \quad & -\left(\sum_{m=0}^{p} S_m \Delta t^{-m}\right)^{-1} \sum_{j=1}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t - j\Delta t) \\ & -\left(\sum_{m=0}^{p} S_m \Delta t^{-m}\right)^{-1} W(t), \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} Y(t) = \quad & \left(I + \sum_{m=0}^{p} S_m \Delta t^{-m}\right) Y(t) \\ & + \sum_{j=1}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t - j\Delta t) + W(t), \end{aligned} \tag{2.9}$$

where $S_p = -I$. ■

Proof. *The proof is presented in Appendix 3.9.3.* □

Since the SVAR and the CVAR models bijectively correspond to each other, Eq.(2.9) corresponds to the SVAR model in Eq.(2.3). Therefore, by comparing Eq.(2.3) and Eq.(2.9), we obtain the following representation of coefficient matrices of the SVAR in Eq.(2.3) by coefficient matrices of the CVAR model in Eq.(2.6).

$$\Psi_0 = I + \sum_{m=0}^{p} S_m \Delta t^{-m}, \tag{2.10}$$

$$\Psi_j = (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m}, \tag{2.11}$$

where $S_p = -I$. Thus, from a given CVAR model, we uniquely derive a SVAR using Eq.(2.10) and (2.11).

Furthermore, we derive the coefficient matrices of the CVAR model, S_m, from the given SVAR model by the following theorem.

Theorem 1. *Under Assumptions 1 and 2, the coefficient matrices of the CVAR model in Eq.(2.6) are represented by coefficient matrices of the SVAR model in*

Eq.(2.3) as follows.

$$S_0 = \Delta t^{-p} I - \sum_{m=1}^{p-1} \left\{ (-1)^m \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + (-1)^{p+m-1} \Delta t^{-p} \frac{p!}{(p-m)!m!} I \right\} + (-1)^p \Delta t^{-p} \Psi_p^{-1} (I - \Psi_0), \quad (2.12)$$

$$S_m = (-1)^m \Delta t^m \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + (-1)^{m+p-1} \Delta t^{m-p} \frac{p!}{(p-m)!m!} I, \quad (2.13)$$

where $1 \leq m \leq p-1$ *and* $S_p = -I$. ■

Proof. *The proof is presented in Appendix 2.7.4.* □

Equations (2.10)-(2.13) indicate a bijective correspondence between the CVAR model and its SVAR model.

We further deduce the SVAR from the DVAR model. As was shown in subsection 2.2.2, we need to know Q or Ψ_0 for this deduction. From Theorem 1, we derive the following Theorem 2, which provides $I - \Psi_0$ from the DVAR matrices only.

Theorem 2. *Under Assumptions 1 and 2, the matrix* $I - \Psi_0$ *is represented by the DVAR parameter matrix as follows.*

$$I - \Psi_0 = (-1)^{p+1} \Delta t^{-p} \Phi_p^{-1}. \quad (2.14)$$

■

Proof. *The proof is presented in Appendix 3.9.5.* □

Thus, Theorem 2 and Eq.(2.4b), (2.5b) indicate a bijective correspondence between the SVAR model and its DVAR model.

From these two theorems, we immediately provide the following corollary on the canonicality of the DVAR model.

Corollary 1. *A DVAR model representing a stable, controllable, continuous time, multivariate, linear Markov system is canonical, and has a bijective correspondence with SVAR and CVAR models of the system.* ■

Here, we summarize all assumptions required by our proposed method. In our study we assume that the objective system has the characteristics of

(1) linear, (2) continuous, (3) Markov, (4) stable, (5) controllable, and
(6) that the approximation error of the backward higher order finite difference for the time discretization of the CVAR model is sufficiently small.

Under these assumptions, the consequences provided above enable us to obtain the SVAR model from the given DVAR model and to derive the CVAR model from the SVAR model. These assumptions hold in the most of scientific and engineering systems in practice and are used in the conventional DVAR modeling. This is one of the main reasons for the wide application of the DVAR modeling. Therefore, our consequences are considered to have a wide applicability similarly to the conventional DVAR modeling. This is a significant advantage of our framework in comparison with the past SVAR modeling methods having strongly limited applicability as explained in subsection 2.2.1.

Based on these considerations, we developed a novel approach, which we call Continuous and Structural Vector AutoRegressive (CSVAR) modeling (Demeshko et al., 2013). The algorithm of the CSVAR modeling approach is shown on Fig. 2.1. It allows us to derive both canonical models, CVAR and SVAR, once we properly derive the DVAR model from a given time series data set observed from a continuous time, multivariate, linear Markov system. In this algorithm, we firstly estimate a DVAR model by applying some traditionally used methods such as the Maximum Likelihood method which computes a likelihood function using a Kalman filter algorithm and applies a quasi-Newton algorithm to search for the maximum of the log-likelihood function (Shea, 1987). Then, we obtain a SVAR model, Eq.(2.3), by deriving its parameter matrices from the matrices of the DVAR model through Eq.(2.4b) and Theorem 2. The derived SVAR model represents the physical dynamics of the system in canonical manner, and the parameter matrices Ψ_j give us the information on the fundamental processes in the system. Finally, by using the relations between the SVAR and the CVAR models presented in Theorem 1, we estimate parameters of the CVAR model of the system (Demeshko et al., 2013).

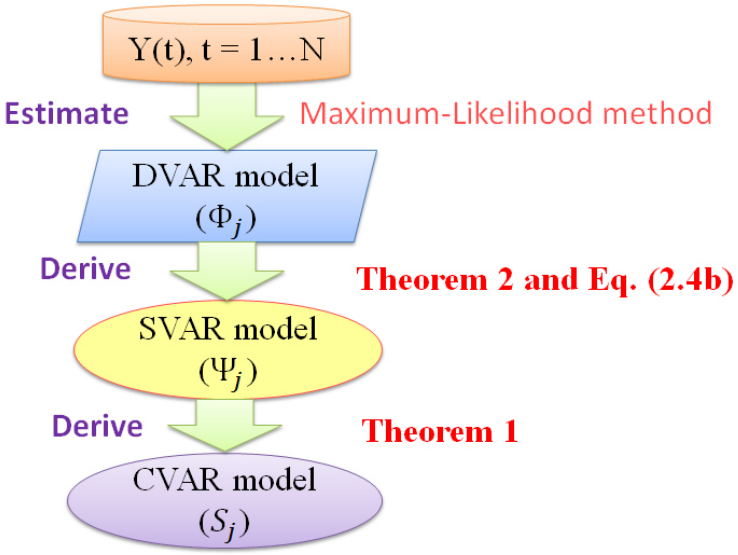

Figure 2.1: The scheme of the CSVAR algorithm.

2.4 Performance on artificial data

In this section, we address objective (2.4) stated in Section 2.1. We demonstrate the performance of the proposed CSVAR modeling through numerical experiments using artificial and real world data. First, we present an illustration of the CSVAR modeling application to a simple physical system. Then we evaluate the accuracies of the canonical SVAR and CVAR models derived from the artificial data by using our modeling method in comparison with their original models that were used to generate the data. Additionally, we confirm the applicability and the accuracy of our CSVAR approach in comparison with a past representative SVAR modeling method. Finally, we evaluate the practicality of CSVAR approach through its application to a real world experimental data.

2.4.1 Illustrative example of CSVAR application

We illustrate the effectiveness of the CSVAR modeling approach through its application to a simple coupled oscillator shown in Fig. 2.2. The dynamics of the system

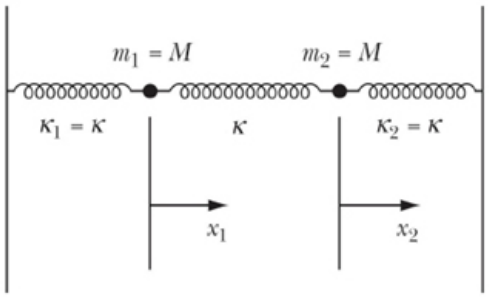

Figure 2.2: Coupled oscillators.

is represented by the following two differential equations.

$$\ddot{x}_1 = \frac{-2\kappa x_1}{M} - \frac{c}{M} v_1, \tag{2.15}$$

$$\ddot{x}_2 = \frac{-2\kappa x_2}{M} - \frac{c}{M} v_2. \tag{2.16}$$

The mass of each object is M, and the spring constant of each spring is κ where their right most and left most ends are fixed at the two walls. Two state variables x_1 and x_2 represent the deviations of the mass positions from their equilibrium, and $v_1 = \dot{x}_1$ and $v_2 = \dot{x}_2$ are their velocities. We also assume some damping forces acting on the masses caused by air friction with coefficient c.

This system is exactly represented by a controllable canonical form of the state space model in continuous time domain shown in Eq. (2.17). This model explicitly indicates kinematics, air friction and observation errors. It consists of the linear differential system equations having external process noises and observation equations of the state variables (Hinrichsen & Pritchard, 2005).

$$\begin{aligned} \frac{dX}{dt} &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-2\kappa}{M} & \frac{\kappa}{M} & -\frac{c}{M} & 0 \\ \frac{\kappa}{M} & \frac{-2\kappa}{M} & 0 & -\frac{c}{M} \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} W \text{ and} \\ Y &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}^T X, \end{aligned}$$

where $X = [x_1, x_2, v_1, v_2]^T$, $W = [w_1, w_2]^T$, and w_1 and w_1 are the external process

noises of x_1 and x_2. The controllable canonical form of the state space model has a direct bijective correspondence to the CVAR model (Hinrichsen & Pritchard, 2005; Brockwell & Davis, 1991). Thus, for the coupled oscillator we have the following CVAR(2) model.

$$Y^{(2)}(t) = \begin{bmatrix} \frac{-2\kappa}{M} & \frac{\kappa}{M} \\ \frac{\kappa}{M} & \frac{-2\kappa}{M} \end{bmatrix} Y^{(0)}(t) + \begin{bmatrix} -\frac{c}{M} & 0 \\ 0 & -\frac{c}{M} \end{bmatrix} Y^{(1)}(t) + W(t). \quad (2.17)$$

We gave the values of the spring constant κ=0.1 N/m, the mass M=1 kg, the air resistance coefficient for the mass $c = 0.5Ns/m$, the period of oscillation $T = 2\pi\sqrt{\frac{M}{\kappa}} = 19.9$ s. By using Eq.(2.10), (2.11) and the relation of Eq.(2.4a), we further transformed this CVAR model to its corresponding SVAR and DVAR models under a time granularity $\Delta t = 1$ s. To ensure the conditions required for the DVAR modeling, we confirmed the stability and the controllability of the obtained DVAR model. By using the derived DVAR model parameters and Eq.(2.1), we further generated a time series $Y(t)$ of 1000 data points. The external bivariate noises $U(t)$ were generated by using an i.i.d. $N(0, \sigma^2)$ distribution where σ is randomly chosen from [0.3,0.7] to maintain the identifiability of the DVAR model.

Finally, we applied the CSVAR algorithm shown in Fig. 2.1 to this sampled time series and estimated the DVAR, SVAR and CVAR models, using provided correct models order, p=2. Table 2.1 shows the comparison between the original models derived from the CVAR model under the discretization with $\Delta t = 1$ s and the estimated models. We see that the SVAR and CVAR models estimated by the CSVAR approach and their original models match well. We also note that the original DVAR model matches well with the DVAR model estimated by Maximum-Likelihood method. The last corresponds to Corollary 1, showing that the objective system is represented by a unique DVAR model when it is linear Markov, stable and controllable system in continuous time domain. Accordingly, we see that in this example the CSVAR modeling appropriately reconstructs the original SVAR and CVAR models of the system from a given time series. These models provide the correct canonical relations between the variables in the original system.

Table 2.1: The parameter matrices of the original and estimated DVAR, SVAR and CVAR models, where Z and $\hat{Z}$ represent original and estimated matrices, respectively, for $Z = \Phi, \Psi$, and S.

DVAR model			SVAR model			CVAR model		
Φ_1	1.476	0.087	Ψ_0	−0.700	0.100	S_0	−0.200	0.100
	0.087	1.476		0.100	−0.700		0.100	−0.200
$\hat{\Phi}_1$	1.469	0.088	$\hat{\Psi}_0$	−0.709	0.105	$\hat{S}_0$	−0.207	0.101
	0.089	1.474		0.103	−0.701		0.102	−0.203
Φ_2	−0.590	−0.035	Ψ_1	2.50	0	S_1	−0.5	0
	−0.035	−0.590		0	2.50		0	−0.5
$\hat{\Phi}_2$	−0.587	−0.036	$\hat{\Psi}_1$	2.502	−0.004	$\hat{S}_1$	−0.502	0.004
	−0.037	−0.590		−0.001	2.498		0.001	−0.498
			Ψ_2	−0.67	0			
				0	−0.67			
			$\hat{\Psi}_2$	−1.00	0			
				0	−1.00			

2.4.2 Accuracy of the proposed method

We demonstrate the accuracy of the SVAR and the CVAR models derived by the CSVAR modeling. For this purpose we perform a set of computer simulations. The procedure of the numerical experiments is similar to the one described in the illustrative example, and it is as follows.

(1) We artificially generate parameter matrices of the CVAR model, $S_m(j = 0, \ldots, p-1)$, each element of which is generated by a uniformly distributed random value in the interval (-1.5,1.5).

(2) Then we generate a CVAR time series data $Y(t)$ by using Eq.(2.8) under a time granularity $\delta t = 0.1\Delta t$ to approximately simulate a continuous process. We also generate a multivariate i.i.d. Gaussian time series W(t). The mean value of each element in W(t) is set to be zero, and its standard deviation is

randomly chosen from [0.3, 0.7] to maintain the identifiability of the CVAR model. We check if this CVAR model satisfy our Assumption 1. If not, we repeat the process from (1) to (2).

(3) We transformed this CVAR model to its corresponding SVAR and DVAR models by using Eq.(2.10), (2.11) and Eq.(2.4a), respectively, under a time granularity Δt. Then, we check the conditions of the stability and controllability of the transformed DVAR model similarly to the illustrative example in Section 4.1. If they are not held, we repeat the process from (1) to (3), till we obtain parameters for the stable and controllable model.

(4) We estimate the AR matrices, $\Phi_j (j = 1, \ldots, p)$, as $\hat{\Phi}_j (j = 1, \ldots, p)$ from the generated multivariate time series $Y(t)$ by a representative DVAR modeling algorithm. Here, we use the Maximum Likelihood method to derive AR parameter matrices (Shea, 1987).

(5) We estimate the SVAR matrices, $\Psi_j (j = 0, \ldots, p)$, as $\hat{\Psi}_j (j = 0, \ldots, p)$ by Eq.(2.4b) and Eq.(2.14). Subsequently, we estimate the CVAR matrices, $S_m (j = 0, \ldots, p-1)$, as $\hat{S}_m (j = 0, \ldots, p-1)$ by Eq.(2.12) and Eq.(2.13).

(6) We evaluate the accuracy of the estimated matrices over the original matrices, using the following cosine measure that represents an element-wise accuracy averaged over all matrices in a model.

$$A_X = \frac{1}{p} \sum_{k=1}^{p} \frac{\sum_{ij} \hat{x}_{k,ij} x_{k,ij}}{\sqrt{\sum_{ij} \hat{x}_{k,ij}^2} \sqrt{\sum_{ij} x_{k,ij}^2}}, \tag{2.18}$$

where $x_{k,ij}$ is the i,j-element of an original matrix X_k which is the AR matrix of the k-th order, and $\hat{x}_{k,ij}$ is the i,j-element of an original matrix $\hat{X}_k$ which is an estimation of X_k. The summation $\sum_{ij}$ is taken over all elements in X_k and $\hat{X}_k$, respectively. Thus, Eq. (2.18) represents an element-wise accuracy averaged over all AR matrices in a model. We apply it to evaluate the accuracy of $X_k = \Phi_j$, Ψ_j or S_m where $k = j$ or m.

We chose a default parameter setting of the dimension of $Y(t)$, $d = 5$, the number of time steps of the data used for modeling of $Y(t), N = 1000$, the order

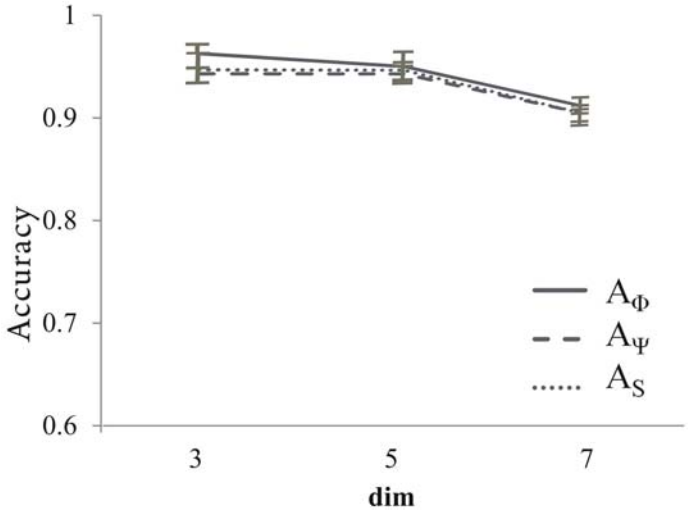

Figure 2.3: Accuracy over different dimensions d, when p=2 and N=1000.

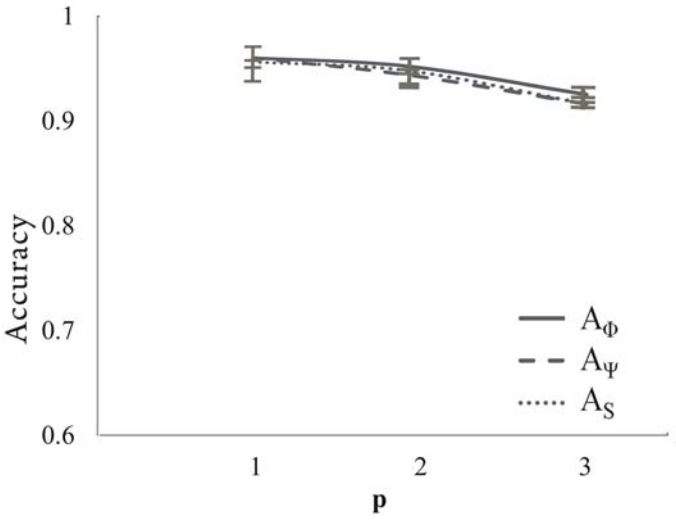

Figure 2.4: Accuracy over different AR orders p, when d=5 and N=1000.

of the CVAR model, $p = 2$, and time granularity $\Delta t = 1$ s, for data generation. Then, we assessed the estimation accuracy over various values of each parameter while setting the other parameters to their default values. For every parameter setting, we repeated 20 experiments and evaluated their 20 accuracies A_X for each experiment.

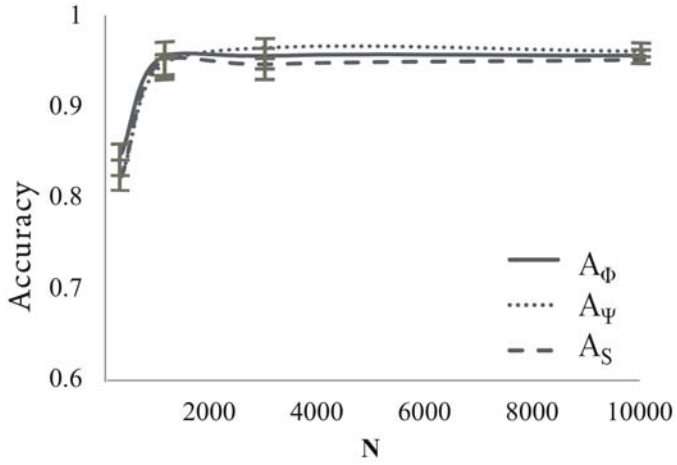

Figure 2.5: Accuracy over different steps N, when d=5 and p=2.

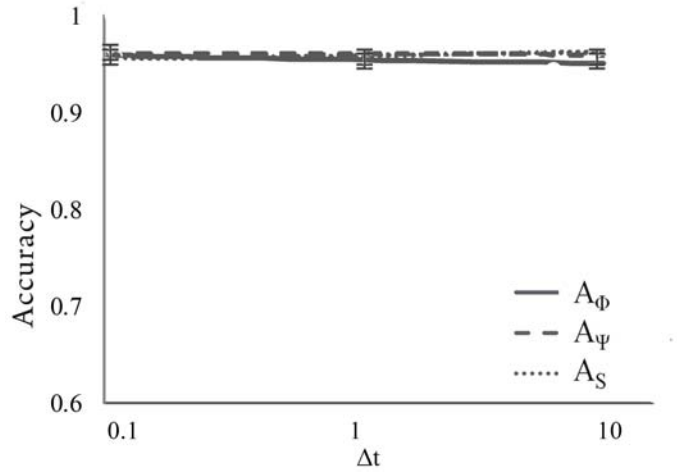

Figure 2.6: Accuracy over different Δt, when d=5, p=2 and N=1000.

Figures 2.3, 2.4, 2.5 and 2.6 show the comparisons of the estimation accuracy over the various values of every parameter. In Fig. 2.3, the lines show the accuracies of AR matrix estimations averaged over the 20 experiments for the DVAR, the SVAR and the CVAR models, respectively, under multiple dimensions of $Y(t)$, $d = 3, 5$ and 7, where the error bars represent the standard deviations. We observe the high accuracies of the SVAR and the CVAR matrices, A_Ψ and A_S, evaluated

by the proposed CSVAR approach as well as the ones of the DVAR matrices, A_Φ, obtained from the data set for all dimensions of the observation vectors. However, for the larger dimension d the accuracy slightly degrades. This was expected, since the number of parameters to be estimated in the AR matrices is $O(d^2)$ and many parameters under a large d make the estimation of the DVAR model statistically unstable for the same length of the given time series data. In Fig. 2.4, we see the results for different AR orders, $p = 1, 2$ and 3. Here, we also note that all three estimated models have high accuracies in all cases. However, similarly to the cases of the large d, the model estimation of the higher orders p shows some degraded accuracies, since the model becomes more complex with more parameters to be estimated. Figure 2.5 shows the results for the data sets with the different number of steps, $N = 300, 1000, 3000$ and 10000. Here, we observe limited accuracy in the small sample case, $N = 300$. This is easily explained by the statistical instability. Under the larger N, the accuracies of all three models estimations significantly high. Finally, in Fig. 2.6, we see the results for different time granularity $\Delta t = 0.1, 1$ and 10. We see that overall results do not depend on parameter Δt as far as the conditions of Assumption 1, the stability and the controllability of the DVAR model hold.

In short summary, our proposed CSVAR modeling accurately captures the system dynamics and the dependency structure among its variables in forms of the SVAR and the CVAR models.

2.4.3 Comparison of CSVAR with AR-LiNGAM

In this subsection, we compare our proposed CSVAR modeling with a past representative method for deriving the SVAR model and evaluate its applicability to various conditions. As was shown in Section 2.2.1, most of the SVAR modeling approaches (Gottschalk, 2001; Pfaff & Kronberg, 2008; Kilian, 2011) require some strong prior knowledge which makes them incomparable with the CSVAR approach. However, AR-LiNGAM method (Hyvarinen et al., 2008) requires weaker assumptions on the system such as acyclic Ψ_0 and non-Gaussian noises. Therefore, we concentrate on the comparisons with AR-LiNGAM. The AR-LiNGAM approach derives the ex-

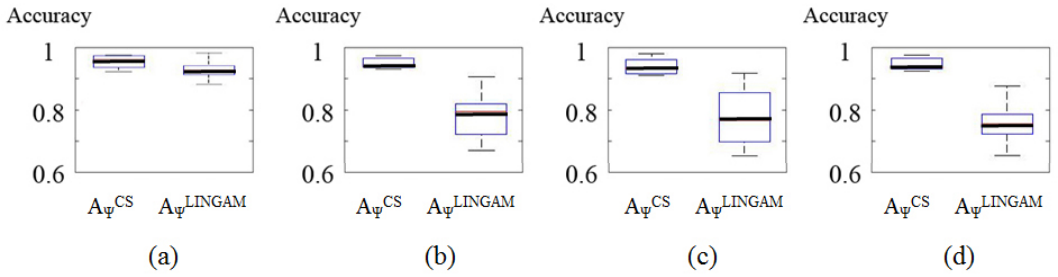

Figure 2.7: Accuracies of SVAR model estimations by CSVAR and AR-LiNGAM for different processes under d=5, N=1000 and p=2.

ternal noise time series, $U(t)$, by the maximum-likelihood estimation of the DVAR model, then it obtains ordering information of variables in Y(t) by applying Independent Component Analysis (ICA). This ordering information together with the orthonormality of the noises further provides matrix Ψ_0 of the SVAR model.

We compared the CSVAR and the AR-LiNGAM approaches by generating four sets of artificial data, (a) non-Gaussian and acyclic case, (b) Gaussian and acyclic case, (c) non-Gaussian case without the acyclicity assumption, and (d) Gaussian case without the acyclicity assumption. To generate non-Gaussian noise, we independently draw the noise values in $W(t)$ from Gaussian distributions and subsequently pass them through a power non-linearity (raising the absolute value to an exponent in the interval [0.5, 0.8] or [1.2, 2.0], but keeping the original sign) to make them non-Gaussian (Shimizu et al., 2011). To generate Gaussian noise, we use the identical process with that of the previous subsection. To generate the acyclic case, we produce SVAR parameter matrices, where matrix Ψ_0 is strictly lower-triangular for the necessary condition of acyclicity. Then, we check regularity of the matrix Φ_p for Eq.(2.12) to ensure the existence of a CVAR model corresponding to the generated SVAR. If the CVAR model exists, we continue the model generation process in the same way described in the previous subsection. To generate the case without the acyclicity assumption, we do not limit Ψ_0 to a strictly lower-triangular matrix.

Thereafter, we apply both our proposed CSVAR and the AR-LiNGAM to the generated data sets. Both approaches give us SVAR model parameter matrices

$\Psi_j^{CS}(j = 0, \ldots, p)$ and $\Psi_j^{LiNGAM}(j = 0, \ldots, p)$. To evaluate the accuracies of the estimated matrices, we compare them with the original ones by Eq.(2.18). In Fig. 2.7(a), we see the box plots of the accuracies of both, CSVAR and the AR-LiNGAM approaches, where a bold line is the median, lower and higher edges of a box are the 25th and 75th percentiles, and the length of a whisker represents the lowest datum within 1.5 IQR (interquartile range) of the lower quartile and the highest datum within 1.5 IQR of the upper quartile. Through these comparisons, we see that both modeling approaches show a very good performance in non-Gaussian and acyclic case. In Fig. 2.7(b), (c) and (d), we see that the accuracies of the AR-LiNGAM are substantially lower than those of our CSVAR method, in the cases when the assumptions of non-Gaussianity and acyclicity are not met. Accordingly, the applicability of the AR-LiNGAM approach is limited to the non-Gaussian and acyclic cases, while our proposed CSVAR modeling is widely applicable to the continuous time linear Markov system as far as the system is stable and controllable.

2.4.4 Performance evaluation by using real world data

In this section, we present the application of the CSVAR approach to a real world experimental data to evaluate its practicality. In our study, we used reactor noise time series measured in an impulse fast neutron research reactor named IBR-2 at Joint Institute of Nuclear Research in Dubna, Russia (Pepyolyshev, 1988). This reactor has a unique structure. It uses rotating main and additional neutron reflectors driven by motors for power pulse initiation. They reflect the generated neutrons back to the core, when they approach to the reactor core. This increases the number of neutrons and activates the fission chain process in the reactor core. That occurs in a very short period and produces power pulses, since both reflectors rotate very fast. Also, according to specifications of the reactor design, the effect of the neutron reflection of the main neutron reflector is almost four times bigger than that of additional neutron reflector.

We analyzed the time series of the peak values of the power pulses, Q, axial deviations of the main neutron reflector, X_Q, and that of the additional neutron reflector, X_A, measured during the stable reactor operation. The axial deviations of

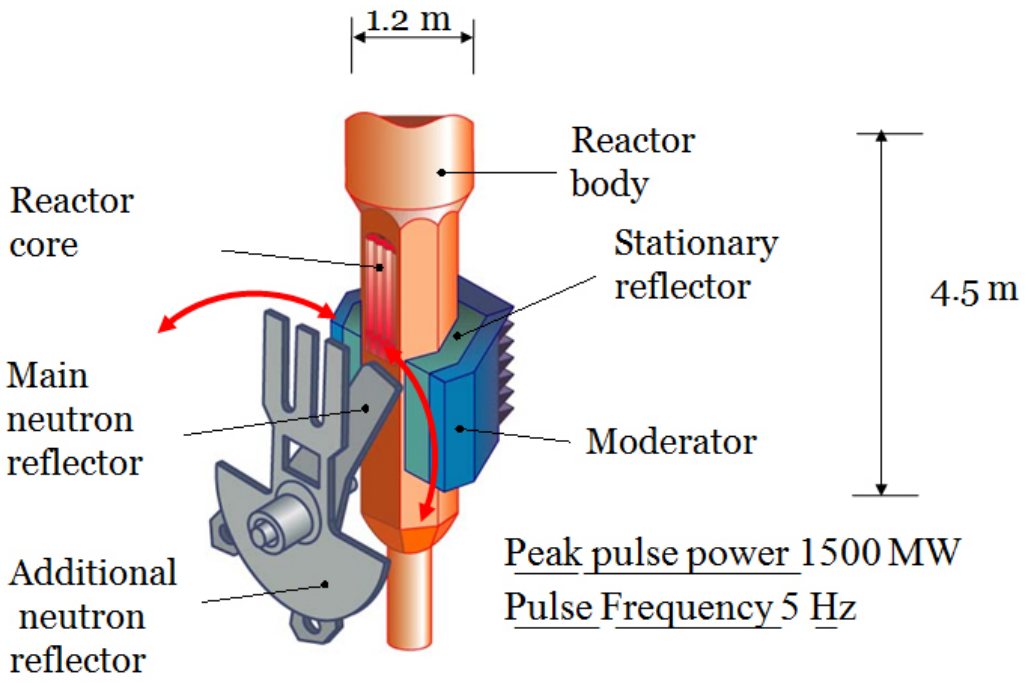

Figure 2.8: The outline of IBR-2 impulse fast neutron reactor.

the reflectors are their angular deviation from the vertical central line of the reactor core. This time series contains 8192 time step measurements. Every variable of the data set has been normalized to give a zero mean and a unit standard deviation. Sampling frequency of the time series data is equal to the frequency of the pulse operation of IBR-2, which is 5 Hz (Pepyolyshev, 1988).

It is known that the relations between the heat removal from the core and the negative feedback effect of core temperature to power generation is approximately represented by the second order delay process. Also the sinusoid impact of periodically rotating reflectors is approximated by the second order delay process. Accordingly, the dynamics of this reactor can be approximated by DVAR(2). After the DVAR(2) estimation by the Maximum Likelihood method, we further calculated its SVAR and CVAR parameters using the CSVAR approach. The results are presented in Eq.(2.19)-(2.25).

Equation (2.19) shows the parameter matrix Ψ_0 of the SVAR model that rep-

resents the fast effects among variables, *i.e.*, the effects which propagate within less than a sampling period. We see the significant values of the $(1,2)$ and$(1,3)$ elements of the matrix, which correspond to the influence of the main neutron reflector's axial deviation, X_Q, on the peak values of the power pulses, Q, and to the influence of X_A on Q, respectively. Also, the ratio of these two numbers is close to four, which coincides with the ratio of both reflectors' efficiency on the neutron reflection. Further, we see that the other elements of Ψ_0 are relatively small, which means that there is no impact from the power output to deviations of both neutron reflectors, and no influence between the reflectors. This result corresponds to the system dynamics of IBR-2 reactor, where both reflectors initiate the power pulse and the rotation of the reflectors is independently driven by motors, *i.e.*, it doesn't depend on the reactor power.

$$\Psi_0 \;=\; \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 1.09 & 11.79 & 2.72 \\ 0.60 & -0.86 & 0.10 \\ -0.53 & 0.53 & -1.11 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array} . \tag{2.19}$$

Equation (2.20) shows the matrix Ψ_1 of the SVAR model which represents the delayed effects among the variables. We see, that Ψ_1 has the structure similar to Ψ_0. This implies that the impacts of the two neutron reflectors are also significant in the first order delay effect. This result is also consistent with the aforementioned physical background of the IBR-2 reactor.

$$\Psi_1 \;=\; \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 3.11 & 90.65 & 18.91 \\ 0.58 & 1.93 & -4.49 \\ 2.80 & -1.51 & -1.78 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array} . \tag{2.20}$$

We note, that Ψ_2 in Eq.(2.21) is diagonal. That well corresponds to Eq.(2.11), from which we have $\Psi_p = (1)^p \Delta t^{-p} S_p$ and $S_p = I$. Therefore, matrix Ψ_2 reflects the fact that we assume the objective system to be a continuous time, linear Markov system

represented by a CVAR process in our CSVAR modeling.

$$\Psi_2 = \begin{array}{c} \begin{matrix} Q & X_Q & X_A \end{matrix} \\ \begin{bmatrix} -25.00 & 0.00 & 0.00 \\ 0.00 & -25.00 & 0.00 \\ 0.00 & 0.00 & -25.00 \end{bmatrix} \end{array} \begin{matrix} Q \\ X_Q \\ X_A \end{matrix} . \tag{2.21}$$

If we look at the DVAR parameter matrices Φ_1 and Φ_2 that represent the delayed effects between the variables, we see that they are not consistent with the dynamics of IBR-2 reactor. Matrix Φ_1 presented in Eq.(2.22) does not show very significant magnitudes of the $(1,2)$ and $(1,3)$ elements in comparison with the other elements. In addition, their values are negative, while the impacts of the neutron reflectors to the power should be positive. The matrix Φ_2 in Eq.(2.23) does not show very clear structure, either. Such inconsistency occurs, since the fast and the delayed effects are not decomposed in the DVAR model.

$$\Phi_1 = \begin{array}{c} \begin{matrix} Q & X_Q & X_A \end{matrix} \\ \begin{bmatrix} -0.021 & -1.689 & -0.492 \\ -0.007 & -0.118 & 0.066 \\ -0.059 & -0.158 & -0.031 \end{bmatrix} \end{array} \begin{matrix} Q \\ X_Q \\ X_A \end{matrix} . \tag{2.22}$$

$$\Phi_2 = \begin{array}{c} \begin{matrix} Q & X_Q & X_A \end{matrix} \\ \begin{bmatrix} 0.476 & -0.654 & -0.203 \\ 0.021 & 0.481 & -0.062 \\ 0.055 & -0.138 & 0.496 \end{bmatrix} \end{array} \begin{matrix} Q \\ X_Q \\ X_A \end{matrix} . \tag{2.23}$$

The proposed CSVAR approach also provides the CVAR model of the nuclear reactor system. This model is canonical and gives us the information on relations among the reactor's processes in continuous time domain. The structure of the CVAR parameter matrices presented in Eq.(2.24) and Eq.(2.25),

$$S_0 = \begin{array}{c} \begin{matrix} Q & X_Q & X_A \end{matrix} \\ \begin{bmatrix} 24.27 & 161.46 & 46.64 \\ -2.21 & 26.60 & 0.57 \\ -3.06 & 4.50 & 21.93 \end{bmatrix} \end{array} \begin{matrix} Q \\ X_Q \\ X_A \end{matrix} . \tag{2.24}$$

$$S_1 = \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 9.38 & -18.13 & -3.78 \\ -0.12 & 9.61 & 0.90 \\ -0.56 & 0.30 & 10.36 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array} . \tag{2.25}$$

is similar to that of the SVAR matrices. Both, S_0 and S_1, have significant values in $(1,2)$ and $(1,3)$ elements, whose ratio is closed to four. However, these elements are negative in S_1, because of the negative feedback of the peak power. This effect occurs as follows. Once the neutron population is increased in the core by the reactor, the core temperature is increased through the activation of the nuclear fission chain reaction. The increase of the temperature reduces the efficiency of the individual nuclear fission reaction, and this suppresses the power generation. These feedback processes are reflected by the negative signs of the main reflectors' impacts in S_1.

In short summary, the application of our CSVAR modeling gives us information on the dependency structure of nuclear reactor processes. Particularly, it presents the fast influences which propagate in the system during a sampling period. The last became possible through the mathematical reconstruction of the SVAR and the CVAR models out of the DVAR model.

2.5 Discussions

The CSVAR modeling approach presented in this study belongs to the framework of canonical modeling of multivariate, linear Markov systems. There are few studies on the canonical and/or causal modeling of the dynamic Markov systems from given time series data in the statistical causal inference (Moneta et al., 2010; Hyvarinen et al., 2008; Gottschalk, 2001; Pfaff & Kronberg, 2008; Kilian, 2011). The advantage of the CSVAR approach is that CSVAR is based on very generic assumption and does not require any specific domain knowledge. In this sense, the CSVAR approach is more comparable with the study (Voortman, Dash, & Druzdzel, 2010), which proposed a method to learn causal structures of the continuous time Markov systems in the framework of the statistical causal inference. Though its basic frame-

work and objective is different from our proposed method, it shares some similar features with ours such that its model consists of higher order time difference variables and requires faithfulness of the objective system ensured by excluding its equilibrium states. The latter feature seems to be associated with the requirement of the controllability in our approach, since some parts of the system can stay at the equilibrium unless the system is fully controllable by the random noises.

This methodology is applicable in many fields, *e.g.*, reactor noise analysis, economics, bioinformatics and so on. The canonical models derived by CSVAR approach can be used to analyze the observed system, even when background knowledge is limited and its inside processes are unknown. The CVAR and the SVAR models provided by this approach bijectively correspond to the objective system dynamics and give us the important information on the system's structural change. Hence, our CSVAR approach enables us to empirically derive scientific models and their associated laws in the system. Such models can be used to monitor and diagnose anomalies of an objective system through time series measurements. For example, in the case of the IBR-2 reactor presented in Section 2.4.4, if we observe some anomalous changes at $(1, 2)$ and $(1, 3)$ elements of the CVAR and the SVAR matrices, we can infer that defects of the neutron reflectors and/or the neutron generation process are occurring.

2.6 Conclusion

In this chapter, we achieved its all four objectives. First, we showed that the DVAR model is canonical upon an assumption on the objective continuous time, multivariate, linear Markov system. Second, we discovered mathematical relations between the CVAR and the SVAR models and the DVAR model of the system. Third, by applying our proposed CSVAR modeling, we accurately derived a canonical representation of an objective system behind a given time series data set under a generic assumption which does not limit its applicability when the system is well approximated by a continuous time, multivariate, linear Markov system. Finally, we demonstrated the applicability and the accuracy of the CSVAR modeling through some numerical experiments using both artificial and real world data, where it

showed a good performance.

We conclude that the CSVAR modeling provides a highly generic methodology to empirically derive a unique model which reflects some elementary rules governing the objective system. Furthermore, the models derived by our approach can help us to improve or discover new knowledge on the system.

However, unless the objective system is described by AR processes, the CVAR and SVAR models derived by CSVAR approach only approximately describe a continuous time linear Markov system. To overcome this issue, in the next chapter, we will develop more advanced modeling approach to derive the CARMA and SARMA models that are exact canonical representations of the system.

2.7 Appendix

2.7.1 The proof of Lemma 2.

Proof. Because of the controllability of the system, the CVAR model of Eq.(2.6) has its controllable canonical form of a state space model (Brockwell & Davis, 1991; Stamer et al., 1996) as follows.

$$\begin{aligned} dX_c(t)/dt &= A_c X_c(t) + B_c W(t), \\ Y(t) &= C_c^T X_c(t), \end{aligned} \tag{2.26}$$

where $X_c(t)$ is a dp-dimensional state variable vector $X_c(t) = [x_c^{(0)}(t)^T \ \ldots \ x_c^{(p-1)}(t)^T]^T$, which is a concatenation of the m-th time derivative of a d-dimensional state variable vector $x_c(t)(= x_c^{(0)}(t))$, $m = 0, \ldots, p-1$. The $dp \times dp$ matrix A_c, $dp \times d$ matrices B_c and C_c are given as follows.

$$A_c = \begin{bmatrix} 0 & I & 0 & \ldots & 0 \\ 0 & 0 & I & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I \\ S_0 & S_1 & S_2 & \ldots & S_{p-1} \end{bmatrix},$$

$$B_c = [\ 0\ 0\ \ldots\ 0\ I\ ]^T \text{ and } C_c = [\ I\ 0\ \ldots\ 0\ ]^T,$$

where $S_m, m = 0, \ldots, p-1$ are given in Eq. (2.6). When the system is stable, all eigenvalues of the system matrix in the state space model, *i.e.*, the solutions z of $det|\sum_{m=0}^{p} S_m z^{-m}| = 0$, have negative real parts (Brockwell & Davis, 1991; Stamer et al., 1996). This implies that $det|\sum_{m=0}^{p} S_m \Delta t^{-m}| \neq 0$, since $\Delta t > 0$. Thus, $\sum_{m=0}^{p} S_m \Delta t^{-m}$ is a regular matrix. □

2.7.2 The proof of Lemma 2

Proof. If the mean of $W^{(2)}(t)$ over time, $E_t[W^{(2)}(t)]$, is nonzero, $W(t)$ does not have the zero mean, and is not stable. This is contradictory to Assumption 1. Thus, $E_t[W^{(2)}(t)] = 0$. Take the second time derivatives of Eq.(2.26) as

$$\begin{aligned} X_c^{(3)}(t) &= A_c X_c^{(2)}(t)/dt + B_c W(t)^{(2)}(t), \\ Y^{(2)}(t) &= C_c^T X_c^{(2)}(t). \end{aligned}$$

$X_c^{(2)}(t)$ is analytically solved as follows.

$$X_c^{(2)}(t) = \exp(A_c t) X_c^{(2)}(0) + \int_0^t \exp(A_c \tau) B_c W^{(2)}(\tau) d\tau.$$

Because this system is stable,

$$\begin{aligned} E_t[Y^{(2)}(t)] &= C_c^T E_t[X_c^{(2)}(t)] = C_c^T E_t\left[\int_0^t \exp(A_c\tau) B_c W^{(2)}(\tau) d\tau\right] \\ &= C_c^T \int_0^t \exp(A_c\tau) B_c E_t[W^{(2)}(\tau)] d\tau = 0. \end{aligned}$$

Similarly, $E_t[Y^{(m+1)}(t)] = 0$ for all $m = 1, \ldots, p$.

The application of the backward higher order finite difference, Eq.(2.7), to each time derivative term in the CVAR model, Eq.(2.6), has the following approximation error $Err(t)$, since its application to $F^{(m)}(t)$ has the error of $O(F^{(m+1)}(t)\Delta t)$.

$$Err(t) = \sum_{m=1}^{p} \alpha_m Y^{(m+1)}(t)\Delta t + \sum_{m=1}^{q} \beta_m W^{(m+1)}(t)\Delta t,$$

where α_m and β_m are nonzero constants. Accordingly, the following holds.

$$E_t[Err(t)] = \sum_{m=1}^{p} \alpha_m E_t[Y^{(m+1)}(t)]\Delta t + \sum_{m=1}^{q} \beta_m E_t[W^{(m+1)}(t)]\Delta t = 0.$$

Thus, the approximation error is unbiased. □

2.7.3 The proof of Lemma 3.

Proof. By substituting Eq.(2.26) into Eq.(2.6), we obtain the following.

$$\frac{1}{\Delta t^p}\sum_{j=0}^{p}(-1)^j\frac{p!}{(p-j)!j!}Y(t-j\Delta t) = \sum_{m=0}^{p-1}S_m\frac{1}{\Delta t^m}\sum_{j=0}^{m}(-1)^j\frac{m!}{(m-j)!j!}Y(t-j\Delta t)+W(t). \tag{2.27}$$

The terms of $Y(t-j\Delta t)$ are summarized with $S_p=-I$ and rewritten by permuting the summations on m and j.

$$\sum_{m=0}^{p}S_m\Delta t^{-m}\sum_{j=0}^{m}(-1)^j\frac{m!}{(m-j)!j!}Y(t-j\Delta t) = \sum_{j=0}^{p}(-1)^j\sum_{m=j}^{p}\frac{m!}{(m-j)!j!}S_m\Delta t^{-m}Y(t-j\Delta t).$$

By substituting these relationships into Eq.(2.27), we obtain the following.

$$-\sum_{m=0}^{p}S_m\Delta t^{-m}Y(t) = \sum_{j=1}^{p}(-1)^j\sum_{m=j}^{p}\frac{m!}{(m-j)!j!}S_m\Delta t^{-m}Y(t-j\Delta t),$$

where $S_p=-I$ and $R_0=I$. By Assumption 1 and Lemma 1, $\sum_{m=0}^{p}S_m\Delta t^{-m}$ is a regular matrix. Then, by multiplying both sides of the equation by $-(\sum_{m=0}^{p}S_m\Delta t^{-m})^{-1}$, we obtain Eq.(2.18). □

2.7.4 The proof of Theorem 1.

Proof. Eq.2.11 becomes as follows in case of $j=p-1$ with $S_p=-I$ defined in Assumption 1.

$$\Psi_{p-1} = (-1)^{p-1}\Delta t^{-p+1}S_{p-1} - (-1)^{p-1}\Delta\mathrm{t}^{\mathrm{-p}}\frac{\mathrm{p!}}{\mathrm{(p\text{-}1)!1!}}\ \mathrm{I}.$$

Therefore,

$$S_{p-1}\Delta t^{-p+1} = (-1)^{p-1}\Psi_{p-1} + \Delta t^{-p}pI. \tag{2.28}$$

Let's assume that Eq.2.13 is an expression for general case $1\le m\le p-1$, we rewrite it as follows.

$$S_m\Delta\mathrm{t}^{\mathrm{-m}} = (-1)^m\sum_{j=m}^{p-1}\frac{\mathrm{j!}}{\mathrm{(j\text{-}m)!m!}}\Psi_j + (-1)^{m+p-1}\Delta\mathrm{t}^{\mathrm{-p}}\frac{\mathrm{p!}}{\mathrm{(p\text{-}m)!m!}}\ \mathrm{I}. \tag{2.29}$$

Note that Eq.2.29 subsumes Eq.2.28. If we write Eq.2.11 for the case $j = p$, we get

$$\Psi_p = (-1)^{p-1}\Delta t^{-p} I. \tag{2.30}$$

Then we rewrite Eq.2.29 as follows.

$$S_m \Delta \mathrm{t}^{\text{-m}} = (-1)^m \sum_{j=m}^{p} \frac{\mathrm{j!}}{\text{(j-m)!m!}} \Psi_j. \tag{2.31}$$

We rewrite Eq.2.11 for the case $j = k-1$ by substituting Eq.2.31 as follows.

$$\Psi_{k-1} = (-1)^{k-1} S_{k-1} \Delta \mathrm{t}^{\text{-(k-1)}} + (-1)^{k-1} \sum_{m=k}^{p} \frac{\mathrm{m!}}{\text{(m-(k-1))!(k-1)!}} (-1)^m \sum_{j=m}^{p} \frac{\mathrm{j!}}{\text{(j-m)!m!}} \Psi_j.$$

By changing the order of the double summation in the last term,

$$(-1)^{k-1} S_{k-1} \Delta \mathrm{t}^{\text{-(k-1)}} = \Psi_{k-1} - (-1)^{k-1} \sum_{j=k}^{p} \Psi_j \sum_{m=k}^{j} (-1)^m \frac{\mathrm{m!}}{\text{(m-(k-1))!(k-1)!}} \frac{\mathrm{j!}}{\text{(j-m)!m!}} I,$$

is obtained. Then, we further obtain the following expression.

$$(-1)^{k-1} S_{k-1} \Delta \mathrm{t}^{\text{-(k-1)}} = \Psi_{k-1} - \sum_{j=k}^{p} \Psi_j \frac{\mathrm{j!}}{\text{(k-1)!(j-k+1)!}} \left(\sum_{u=0}^{j-k+1} (-1)^u \frac{\text{(j-k+1)!}}{(u)!\text{(j-u-k+1)!}} I - I \right),$$

where $m = u + k - 1$. The summation over u is zero based on binomial theorem, and further rewriting $k-1$ in the formula by $m-1$, we know that Eq.2.31 holds for $m-1$. By induction, Eq.2.31 holds for $1 \leq m \leq p-1$. Furthermore, by substituting Eq.2.30 into Eq.2.31, we obtain Eq.2.29 and thus Eq.2.13 for $1 \leq m \leq p-1$. To obtain Eq.2.12, we substitute $S_p = -I$ into Eq.2.10.

$$I - \Psi_0 = -S_0 - \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} + \Delta t^{-p} I.$$

By substituting Eq.2.30 into this equation, we obtain next formula.

$$I - \Psi_0 = (-1)^{p-1} \Delta t^p \Psi_p \left(\Delta t^{-p} I - \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} - S_0 \right). \tag{2.32}$$

From Eq.2.30, Ψ_p is regular. Then, we reformulate Eq.2.32 as follows,

$$\Delta t^{-p} I - \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} - S_0 = (-1)^{p-1} \Delta t^{-p} \Psi_p^{-1} \left(I - \Psi_0 \right).$$

Thus,

$$S_0 = -\sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} + \Delta t^{-p} I + (-1)^p \Delta t^{-p} \Psi_p^{-1} \left(I - \Psi_0 \right).$$

Substituting Eq.2.29 inside the second term of this formula, we obtain Eq.2.12.

□

2.7.5 The proof of Theorem 2.

Proof. By substituting Eq.2.29 into $\sum_{m=1}^{p-1} S_m \Delta t^{-m}$, we obtain the following relation.

$$\sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} = \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} (-1)^m \sum_{j=m}^{p-1} \frac{\mathrm{j!}}{\mathrm{(j\text{-}m)!m!}} \Psi_j + \Delta t^{-p} \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} (-1)^{p+m-1} \frac{\mathrm{p!}}{\mathrm{(p\text{-}m)!m!}}\ \mathrm{I}. \tag{2.33}$$

To derive Ψ_0, we rewrite Eq.2.10 by substituting $S_p = -I$ in Assumption 1 as follows:

$$\Psi_0 = I - \Delta t^{-p} I + \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} S_m \Delta t^{-m} + S_0. \tag{2.34}$$

By substituting Eq.2.12 and Eq.2.33 into this equation, we obtain the following.

$$\begin{aligned}
\Psi_0 = {} & I - \Delta t^{-p} I + \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} (-1)^m \sum_{j=m}^{p-1} \frac{\mathrm{j!}}{\mathrm{(j\text{-}m)!m!}} \Psi_j + \Delta \mathrm{t^{\text{-}p}} \sum_{\mathrm{m=1}}^{\mathrm{p\text{-}1}} (-1)^{p+m-1} \frac{\mathrm{p!}}{\mathrm{(p\text{-}m)!m!}}\ \mathrm{I} \\
& + \Delta t^{-p} I - \sum_{m=1}^{p-1} \left\{ (-1)^m \sum_{j=m}^{p-1} \frac{\mathrm{j!}}{\mathrm{(j\text{-}m)!m!}} \Psi_j + (-1)^{p+m-1} \Delta \mathrm{t^{\text{-}p}} \frac{\mathrm{p!}}{\mathrm{(p\text{-}m)!m!}}\ \mathrm{I} \right\} \\
& + (-1)^p \Delta \mathrm{t^{\text{-}p}} \Psi_{\mathrm{p}}^{\mathrm{\text{-}1}} (\mathrm{I} - \Psi_0) = I + (-1)^p \Delta \mathrm{t^{\text{-}p}} \Psi_{\mathrm{p}}^{\mathrm{\text{-}1}} (\mathrm{I} - \Psi_0).
\end{aligned} \tag{2.35}$$

From Eq.2.4a, we see that $\Phi_p = (I - \Psi_0)^{-1} \Psi_p$, where $I - \Psi_0$ is always regular by Assumption 1 and Eq.2.10. Since Ψ_p is regular by Eq.2.30, Φ_p is also regular. Thus, we write as $\Phi_p^{-1} = \Psi_p^{-1}(I - \Psi_0)$. By substituting it into Eq.2.35, we derive the following expression.

$$\Psi_0 = I + (-1)^p \Delta t^{-p} \Phi_p^{-1}, \tag{2.36}$$

and thus we obtain Eq.2.14. □

Chapter 3

CSARMA modeling approach

3.1 Introduction

As mentioned in previous chapter, many real world systems observed in the discrete time domain are often analyzed by the DVAR model (Brockwell & Davis, 1991). However, in many cases, the DVAR models with finite orders are approximations of discrete time, multivariate, linear Markov systems, while the multivariate discrete time Autoregressive Moving-Average (DARMA) models are their exact notions. Therefore, if the objective system that has produced a given data set is well approximated by a multivariate linear Markov process, we should identify its structure and parameters using DARMA modeling. Accordingly, the DARMA modeling has been widely used in many analyses, *e.g.*, biology (Perrott & Cohen, 1996), medicine (Gannabathula, 1988), finance (Shittu & Yaya, 2009), economy (Chen et al., 2010), physics (Tran, 2003), engineering (Zhao et al., 2007) and energy (Zhang & Xu, 2010). However, the DARMA model is not canonical, and does not have a bijective correspondence with the system dynamics (Gottschalk, 2001; Moneta et al., 2010; Kawahara et al., 2011). This limitation makes it difficult to identify the structure and parameters of the objective system.

As a remedy to this difficulty, a structural ARMA (SARMA) model has been studied. It is known to have a bijective correspondence with the dynamics of a multivariate linear Markov system in a discrete time domain (Gottschalk, 2001; Moneta et al., 2010; Kawahara et al., 2011). Since the estimation of the DVAR model is much easier than of the DARMA model, most studies in this area have concentrated on the SVAR models, which belong to a subclass of the SARMA models. Only a small number of studies have investigated SARMA modeling, but they

require some strong assumptions for model identification, such as acyclicity of an objective system and non-Gaussianity of its external noise (Kawahara et al., 2011), and system representation by a small number of parameters and partial independence of the noise (Mainassara & Francq, 2009). However, by introducing these assumptions and focusing on some special systems, we have significantly limited the applicability of the SARMA models.

On the other hand, linear differential equations provide a standard representation of a generic real world system, which can be well approximated by a continuous time, multivariate, linear Markov process. A CARMA model consists of such higher order differential equations, and is known to bijectively correspond to the system dynamics (Brockwell & Davis, 1991; Stamer et al., 1996). Some past studies characterized the relationship between the CARMA and the DARMA processes under an equivalence measure that the latter generates the time series distribution identical with the former's (Priestley, 1981; Soderstrom, 1991; Brockwell & Peter, 1995; Brockwell & Brockwell, 1999). Particularly, the time series distribution generated by a given stable CARMA(p,q) ($q < p$) process is known to be also generated by a DARMA(p,$p-1$) process (Priestley, 1981). Based on this equivalence, many researchers studied methods to estimate a CARMA model under time series data observed from a continuous time linear Markov process in a discrete time domain (Soderstrom, 1991; Larsson et al., 2006; Mahata & Fu, 2007; Tomasson, 2011; Chambers & Thornton, 2012). However, any closed-form expression to relate the CARMA(p,q) and the DARMA(p,$p-1$) models have not been presented under this equivalence measure. Moreover, the uniqueness of the CARMA(p,q) process generating the given time series data sampled at the discrete time steps and thus the uniqueness of the CARMA(p,q) model corresponding to the DARMA(p,$p-1$) model of the time series data have not been generally clarified yet in these studies. This situation does not necessarily satisfy the needs of scientific and engineering fields including the aforementioned analyses. The identification and the understanding of the objective system are difficult under the ambiguity of the CARMA model for the data given in the discrete time domain. In addition, scientists and engineers cannot sufficiently use their background knowledge across the continuous and discrete time domains without knowing some explicit and simple expression relating

the CARMA and the DARMA models. Particularly, this is required in many cases that they want to understand the systems based on the system structures appearing in the CARMA and the SARMA models.

To address these issues, we target the following four objectives in this chapter.

(3.1) Using some generic assumptions of an objective system and introducing some generic approximations to its modeling, we show that the DARMA model of the system has a bijective correspondence with the system dynamics.

(3.2) Under the assumptions and the approximations in (3.3), we clarify the mathematical bijective relationships in closed-form expressions between the CARMA model, the SARMA model and the DARMA model of the system.

(3.3) We further propose a new modeling approach to derive the CARMA and the SARMA models from the DARMA model that has been estimated using time series data observed from the system and given orders (p,q) of the system.

(3.4) We demonstrate the performance of our proposed approach through some numerical experiments using both artificial and real world data.

Our analysis shows that the DARMA model having its correct system orders (p,q) bijectively corresponds to the dynamics of the stable, continuous time, multivariate, linear Markov system under its controllability and observability across continuous and discrete time domains, and the first order finite difference approximation. We also show that the approximation ensures the statistically unbiased estimator with its consistency under an appropriately small sampling interval. The stability, the controllability and the observability of the system with some fine sampling steps are also required by the DARMA modeling of the continuous time, multivariate, linear Markov system (Brockwell & Davis, 1991; Soderstrom, 1991). Accordingly, the CARMA and the SARMA models can be reconstructed from the DARMA model without imposing strong additional limitations.

This chapter is structured as follows. Section 3.2 discusses some past studies related to our work, and clarifies some technical issues that must be addressed. A discussion on the relations of the ARMA models, assumptions required for our

modeling, and their characterizations are provided in Section 3.3. Our theoretical analyses and proposed approach that address the objectives (3.1), (3.2) and (3.3) are described in Section 3.4. A basic performance evaluation of our proposed approach, its comparison with an existing SARMA modeling technique, and a practical demonstration using a real world application are shown in Sections 3.5 and 3.6 for the objective (3.4).

3.2 Related work

Similar to the section 2.2.1, here, we first explain the SARMA model and the limitation of its past modeling approaches. Then, we further explain the CARMA model together with the finite difference approximation to transform it into a discrete time domain.

3.2.1 SARMA modeling

Given a stable, multivariate, liner Markov system that is controllable and observable in a discrete time domain, take $Y(t)$ to be a d-dimensional variable vector observed from the system at a time, t. Then, we can represent the system using the DARMA(p,q) model with a discrete time interval, Δt (Brockwell & Davis, 1991),

$$Y(t) = \sum_{j=1}^{p} \Phi_j Y(t - j\Delta t) + U(t) + \sum_{j=1}^{q} \Theta_j U(t - j\Delta t). \tag{3.1}$$

$U(t)$ is a d-dimensional stable noise vector having a zero mean and a full rank covariance, and is unobserved and i.i.d. in a discrete time domain. Φ_j, $j = 1, \ldots, p$ and Θ_j, $j = 1, \ldots, q$ ($q < p$) are the respective $d \times d$ AR and MA coefficient matrices. This model contains some ambiguity so that it can represent infinitely many systems, since Eq.(3.1) is equivalent to

$$Y(t) = \sum_{j=1}^{p} \Phi_j Y(t - j\Delta t) + QV(t) + \sum_{j=1}^{q} \Theta_j QV(t - j\Delta t),$$

where Q is a $d \times d$ regular matrix and $V(t) = Q^{-1}U(t)$. This formula can be easily used to show that $Y(t)$ has infinitely various impulse responses for $V(t)$, depending

on the choice of Q (Gottschalk, 2001; Moneta et al., 2010; Kawahara et al., 2011). In other words, for a given observed time series, $Y(t)$, the infinite number of systems that could generate $Y(t)$ are represented by a unique DARMA model. Thus, we need to know a unique combination of Q and $V(t)$ that bijectively corresponds to the objective system dynamics, so that we can identify it. This formula can be rewritten

$$Y(t) = \sum_{j=0}^{p} \Psi_j Y(t - j\Delta t) + V(t) + \sum_{j=1}^{q} \Omega_j V(t - j\Delta t), \tag{3.2}$$

where $\Psi_0 = I - Q^{-1}$, $\Psi_j = Q^{-1}\Phi_j$, and $\Omega_j = Q^{-1}\Theta_j Q$ (Kawahara et al., 2011). Using a unique combination of Q and $V(t)$, Eq.(3.2) is called a structural ARMA (SARMA) model (Gottschalk, 2001; Moneta et al., 2010; Kawahara et al., 2011).

This formula shows that the matrix Q is reflected in a direct dependency structure Ψ_0 among the observed variables in $Y(t)$, and is also reflected in higher order terms in both the AR and MA parts. However, Q is not reproduced from the DARMA model, because it is integrated in Φ_j and Θ_j in indecomposable forms as $\Phi_j = Q\Psi_j$ and $\Theta_j = Q\Omega_j Q^{-1}$. Accordingly, we need to provide Q using extra information to uniquely derive the SARMA model from the DARMA model. A natural idea to choose Q is to make the noise in $V(t)$ mutually uncorrelated by introducing a constraint $QQ^T = E_t(U(t)U(t)^T)$. However, this choice of Q is not unique. For example, for any orthonormal matrix O, we have $QO(QO)^T = QOO^TQ^T = QQ^T$. Then, if Q is a solution, so is QO (Moneta et al., 2011).

As previously discussed, many past studies on the identification of structural models are limited to structural vector autoregressive (SVAR) models that omit the last MA term in Eq.(3.2) (Gottschalk, 2001; Kilian, 2011; Hyvarinen et al., 2008; Lacerda et al., 2008). However, a more generic SARMA model has not been extensively researched. (Mainassara & Francq, 2009) assumed the partial independence of the external noises, and used the coefficient matrices of the SARMA model represented by a small number of parameters based on our domain knowledge. (Kawahara et al., 2011) proposed an approach for SARMA modeling using such assumptions as the acyclicity of the direct dependency among the variables in $Y(t)$, and the non-Gaussianity of external noises in $V(t)$, in addition to the typical assumption that all noises in $V(t)$ are mutually and temporally independent.

Most of the assumptions and the constraints used by these past approaches require much domain knowledge of the objective system. This requirement significantly limits the applicability of these approaches, because such domain knowledge is not readily available in many real-world applications.

3.2.2 CARMA modeling and finite difference method

Given a stable, multivariate, linear Markov system that is controllable and observable in a continuous time domain, it can be represented by the following continuous time ARMA (CARMA) model, consisting of continuous time higher order differential equations with noise terms (Brockwell & Davis, 1991; Stamer et al., 1996; Chambers & Thornton, 2012).

$$Y^{(p)}(t) = \sum_{m=0}^{p-1} S_m Y^{(m)}(t) + W^{(0)}(t) + \sum_{m=1}^{q} R_m W^{(m)}(t), \tag{3.3}$$

where $W(t)$ is a d-dimensional stable noise vector having a zero mean and a full rank covariance, and is unobserved and i.i.d. in a continuous time domain[1]. $Y^{(m)}(t)$ and $W^{(m)}(t)$ are the m-th time derivatives of $Y(t)$ $(= Y^{(0)}(t))$ and $W(t)$ $(= W^{(0)}(t))$. S_m, $m = 0, \ldots, p-1$ and R_m, $m = 1, \ldots, q$ $(q < p)$ are $d \times d$ AR and MA coefficient matrices. This CARMA model is known to have a bijective relation with a controllable canonical form of a linear state space model which is also bijective with the system dynamics (Brockwell & Davis, 1991; Stamer et al., 1996). In other words, a CARMA model bijectively corresponds to the system dynamics. Hence, estimation of the CARMA model from a time series given in a discrete time domain and derivation of the CARMA model from a given DARMA model will be powerful techniques to analyze the dynamics underlying the data.

As mentioned in Section 3.1, the time series distribution generated by a stable CARMA(p,q) $(q < p)$ process is also generated by a DARMA(p,$p-1$) process (Priestley, 1981). But, the uniqueness of the CARMA(p,q) model correspond-

[1] $W(t)$ is equal to $\sigma dE(t)/dt$ in the standard representation of the CARMA model, where $E(t)$ and σ are a normalized external noise vector and its scaling factor, respectively. We use $W(t)$ for its consistency with the noise vectors in the DARMA model as shown in (Chambers & Thornton, 2012).

ing to a given DARMA(p,$p-1$) model of the continuous time system have not been clarified under this equivalence measure[2], and any closed-form expression to relate the two models have not been presented. As an efficient remedy to this limitation, we consider to use finite difference approximation to explicitly relate the CARMA model with the DARMA model.

Here, we apply backward Euler formula upto the m-th derivative shown in Eq. 2.7, which is known to be consistent (Levy & Lessman, 1992; LeVeque, 2007). Therefore, the CARMA model of a stable, continuous time, linear Markov system having its given orders (p,q) can be directly transformed to a DARMA model with its convergent error using a consistent finite difference scheme.

However, to our best knowledge, the generic relationships between the CARMA, the SARMA and the DARMA models have not been extensively investigated based on this approximation.

3.3 Preliminary discussion and analysis

In this section, we first discuss the relationships between the CARMA, the SARMA and the DARMA models. Second, we state our assumptions, and characterize them for our CARMA and SARMA modeling. These provide bases for our detailed analysis and proposal in Section 3.4.

3.3.1 Relationships of ARMA models

A stable, multivariate linear Markov system, which is controllable and observable in both continuous and discrete time domains, can be represented by the SARMA model, Eq.(3.2), and the CARMA model, Eq.(3.3). These are known to bijectively correspond to the system dynamics, respectively, as explained in Section 3.2. Accordingly, the SARMA and the CARMA models have a bijective correspondence.

[2]Many studies investigated the embedding problem of an arbitrary given DARMA model. It is to assess the existence of a CARMA model generating time series sharing an identical distribution with that generated by the DARMA model (Brockwell & Peter, 1995; Brockwell & Brockwell, 1999). Our issue here is not this problem, since we always limit the DARMA model to the one having some corresponding CARMA model.

To further assess this relationship, we represent the DARMA model, Eq.(3.1), and the CARMA model, Eq.(3.3), in the controllable canonical forms of their state space models (Brockwell & Davis, 1991). The DARMA model is represented as

$$\begin{aligned} X_d(t) &= A_d X_d(t-\Delta t) + B_d U(t), \\ Y(t) &= C_d^T X_d(t), \end{aligned} \tag{3.4}$$

where $X_d(t)$ is a dp-dimensional state variable vector $X_d(t) = [x_d(t-(p-1)\Delta t)^T \ \ldots \ x_d(t)^T]^T$, which is a concatenation of the $(j-1)$-th time delay of a d-dimensional state variable vector $x_d(t)$, $j = 1, \ldots, p$. The $dp \times dp$ matrix A_d, $dp \times d$ matrices B_d and C_d are given as follows.

$$A_d = \begin{bmatrix} 0 & I & 0 & \ldots & 0 \\ 0 & 0 & I & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I \\ \Phi_p & \Phi_{p-1} & \Phi_{p-3} & \ldots & \Phi_1 \end{bmatrix},$$

$$B_d = [\ 0 \ \ldots \ 0 \ I \ ]^T \text{ and } C_d = [\ 0 \ \ldots \ 0 \ \Theta_q \ \ldots \ \Theta_1 \ I \ ]^T,$$

where $\Phi_j, j = 1, \ldots, p$ and $\Theta_j, j = 1, \ldots, q$ $(q < p)$ are given in Eq. (3.1). Similarly, the CARMA model is represented as

$$\begin{aligned} dX_c(t)/dt &= A_c X_c(t) + B_c W(t), \\ Y(t) &= C_c^T X_c(t), \end{aligned} \tag{3.5}$$

where $X_c(t)$ is a dp-dimensional state variable vector $X_c(t) = [x_c^{(0)}(t)^T \ \ldots \ x_c^{(p-1)}(t)^T]^T$, which is a concatenation of the m-th time derivative of a d-dimensional state variable vector $x_c(t)(= x_c^{(0)}(t))$, $m = 0, \ldots, p-1$. The $dp \times dp$ matrix A_c, $dp \times d$ matrices B_c and C_c are given as follows.

$$A_c = \begin{bmatrix} 0 & I & 0 & \ldots & 0 \\ 0 & 0 & I & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & I \\ S_0 & S_1 & S_2 & \ldots & S_{p-1} \end{bmatrix},$$

$$B_c = [\ 0\ 0\ \ldots\ 0\ I\ ]^T \text{ and } C_c = [\ I\ R_1\ R_2\ \ldots\ R_q\ 0\ \ldots\ 0\ ]^T,$$

where $S_m, m = 0, \ldots, p-1$ and $R_m, m = 1, \ldots, q$ $(q < p)$ are given in Eq. (3.3). Assuming that the change of $W(t)$ for Δt is negligible, we analytically obtain the following discrete time state space model of the system from Eq.(3.5) (Brockwell & Davis, 1991).

$$\begin{aligned} X_c(t) &= \exp(A\Delta t)X_c(t-\Delta t) + A_c^{-1}(\exp(A_c\Delta t) - I)B_cW(t), \qquad (3.6) \\ Y(t) &= C_c^T X_c(t), \end{aligned}$$

where $\exp(A\Delta t) = I + A\Delta t + A^2\Delta t^2/2 + \cdots = \sum_{k=0}^{\infty} A^k\Delta t^k/k!$. We provide the following lemma using these state space representations.

Lemma 4. *If the change of $W(t)$ over the period Δt is negligible, there exists a $d \times d$ regular matrix P to relate $U(t)$ in Eq.(3.1) and $W(t)$ in Eq.(3.3) as $U(t) = PW(t)$.*

■

Proof. *Because the CARMA and the DARMA models represent an identical system, the two discrete time state space models, Eq.(3.4) and Eq.(3.6), also represent the identical system under the negligibly slow change of $W(t)$ over Δt. From the stability of the system, all eigenvalues of A_c have finite negative real parts. Hence, $\exp(A_c\Delta t)$ and its corresponding A_d are regular. This implies the existence of a $dp \times dp$ regular matrix T to transform Eq.(3.6) to Eq.(3.4) as*

$$\begin{aligned} &X_d(t) = TX_c(t),\ A_d = T\exp(A_c\Delta t)T^{-1}, C_d = C_cT^{-1}, \text{and} \\ &B_dU(t) = TA_c^{-1}(\exp(A_c\Delta t) - I)B_cW(t). \end{aligned}$$

By the definitions of B_c and B_d, the last formula shows

$$TA_c^{-1}(\exp(A_c\Delta t) - I) = \begin{bmatrix} * & \cdots & * & 0 \\ \vdots & \ddots & * & \vdots \\ * & \cdots & * & 0 \\ * & \cdots & * & P \end{bmatrix},$$

where $*$ and 0 are arbitrary $d \times d$ blocks and $d \times d$ zero blocks, respectively. P is a $d \times d$ matrix to relate $U(t)$ and $W(t)$ as $U(t) = PW(t)$. Because of the regularity of A_c and T, $TA_c^{-1}(\exp(A_c\Delta t) - I)$ is also regular under $\Delta t > 0$. This implies the

diagonal block P is a regular matrix, since all elements upper than P are zeros in $TA_c^{-1}(\exp(A_c\Delta t) - I)$. □

Lemma 4 indicates that P is defined by the coefficient matrices of the CARMA and the DARMA models but not the statistical features of $U(t)$ and $W(t)$. As we will see in Section 3.4, such a regular matrix P uniquly exists under the first order finite difference approximation. Here, we take P as the matrix Q to match $V(t)$ in Eq.(3.2) to $W(t)$ in Eq.(3.3) and to make the SARMA model represent the system identical with the CARMA model. In the rest of this chapter, we represent Q and $V(t)$ in the SARMA model by P and $W(t)$, respectively.

As discussed in Section 3.2.1, the matrices Φ_j and Θ_j in the DARMA model, Eq.(3.1), are derived from the matrices Ψ_j and Ω_j in the SARMA model, Eq.(3.2), using P such that $\Phi_j = P\Psi_j$ and $\Theta_j = P\Omega_j P^{-1}$. Also, P has the unique relation $P = (I - \Psi_0)^{-1}$ with Ψ_0. Therefore, these matrices have the following relations (Kawahara et al., 2011).

$$\Phi_j = (I - \Psi_0)^{-1}\Psi_j, \tag{3.7a}$$

$$\Psi_j = (I - \Psi_0)\Phi_j, \tag{3.7b}$$

$$\Theta_j = (I - \Psi_0)^{-1}\Omega_j(I - \Psi_0), \tag{3.8a}$$

$$\Omega_j = (I - \Psi_0)\Theta_j(I - \Psi_0)^{-1}, \tag{3.8b}$$

$$U(t) = (I - \Psi_0)^{-1}W(t), \tag{3.9a}$$

$$W(t) = (I - \Psi_0)U(t). \tag{3.9b}$$

These formulae characterize the relationship between the DARMA model and the SARMA model as follows.

Lemma 5. *There exist infinitely many SARMA models that deduce a given DARMA model. However, there is a unique combination of a DARMA model and a regular matrix P that a given SARMA model deduces.* ■

Proof. *A given combination of $\Phi_j, j = 1, \ldots, p$ and $\Theta_j, j = 1, \ldots, q$, can be deduced using various combinations of $\Psi_j, j = 1, \ldots, p$, $\Omega_j, j = 1, \ldots, q$ and a regular matrix $P = (I - \Psi_0)^{-1}$ through Eq.(3.7a) and (3.8a). Thus, there exist infinitely many*

SARMA models that deduce the given DARMA model. On the other hand, given a combination of Ψ_0, $\Psi_j, j = 1, \ldots, p$ *and* $\Omega_j, j = 1, \ldots, q$, P *is uniquely provided by* $P = (I - \Psi_0)^{-1}$, *and a unique combination of* $\Phi_j, j = 1, \ldots, p$ *and* $\Theta_j, j = 1, \ldots, q$ *is provided through Eq.(3.7a) and (3.8a). Thus, a unique combination of the DARMA model and the regular matrix* P *is deduced from a given SARMA model.* □

As we mentioned earlier, the SARMA and the CARMA models of an objective system have a bijective relation. In contrast, Lemma 5 indicates that the correspondence between the SARMA and the DARMA models is surjective. However, in Section 3.4, we will show that their correspondence is also bijective under the first order finite difference approximation and the aforementioned selection of P.

3.3.2 Assumptions for modeling and their characterization

In this subsection, we explicitly state our assumptions and provide their characterization required by our modeling principles and algorithm. The following is an explicit restatement of our assumption used in the previous sections.

Assumption 3. *Our objective system is a stable, continuous time, multivariate, linear Markov system that is controllable and observable in both continuous and discrete time domains. It has stable external noise components that have a zero mean and a full rank covariance and that is i.i.d. over time.* ■

Under this assumption, we obtain important properties of the CARMA model and the sampling interval Δt as shown in the following two lemmas.

Lemma 6. *Given the CARMA model, Eq.(3.3), with a positive real constant* $\Delta t > 0$, *that represents the objective system under Assumtion 3,* $\sum_{m=0}^{p} S_m \Delta t^{-m}$ *is a regular matrix where* $S_p = -I$. ■

Proof. *The same as for Lemma 1.* □

Lemma 7. *Given the CARMA model, Eq.(3.3), with a positive real constant* $\Delta t > 0$, *that represents the objective system under Assumtion 3,* $\sum_{m=0}^{q} R_m \Delta t^{-m}$ *is a regular matrix where* $R_0 = I$. ■

Proof. *The proof is presented in Appendix 3.9.1.* □

Under Assumption 3, the use of Eq.(2.7) for the first order finite difference approximation ensures the unbiased estimator of the CARMA modeling by the following lemma.

Lemma 8. *The finite difference approximation of the CARMA model using Eq.(2.7) is unbiased under Assumption 3.* ■

Proof. *Is the same as for Lemma 2.* □

Our proposed approach further requires the following second assumption.

Assumption 4. *Let λ_j, λ_{Rj}, λ_{Ij} be an eigenvalue of the matrix A_c in Eq.(3.5), its real part, and its imaginary part, respectively, for $j = 1, \ldots, d$. Then, define the following maximum values,*

$$|\lambda|_{max} = \max_{j=1,\ldots,d} \sqrt{\lambda_{Rj}^2 + \lambda_{Ij}^2} \text{ and } |\lambda_I|_{max} = \max_{j=1,\ldots,d} |\lambda_{Ij}|.$$

In addition, let ω be the angle frequency of the highest frequency component in $W(t)$. Then, we assume that the sampling interval $\Delta t > 0$ satisfies the following three conditions.

$$\text{(a) } \Delta t < 1/|\lambda|_{max}, \text{ (b) } \Delta t < \pi/|\lambda_I|_{max} \text{ and (c) } \Delta t < \pi/\omega.$$

■

The condition (a) $\Delta t < 1/|\lambda|_{max}$ is to ensure the sufficient consistency of our CARMA and the SARMA modeling by the following lemma.

Lemma 9. *If we apply the finite difference approximation of Euler formula to the controllable canonical form of a given continuous time state space model, Eq.(3.5), the approximation errors of the coefficients in the first and the second terms of its l.h.s. are respectively bounded by*

$$\exp(|\lambda|_{max}\Delta t) - 1 - |\lambda|_{max}\Delta t \text{ and } \frac{\exp(|\lambda|_{max}\Delta t) - 1 - |\lambda|_{max}\Delta t}{|\lambda|_{max}}.$$

■

Proof. *The proof is presented in Appendix 3.9.2.* □

The error bound of the coefficients relative to $exp(|\lambda|_{max}\Delta t)$ in the first term is 26.4%, if $\Delta t \simeq 1/|\lambda|_{max}$ following the first constraint. That relative to $(exp(|\lambda|_{max}\Delta t)-1)/|\lambda|_{max}$ in the second term is 41.8%. If we choose smaller Δt, the error bounds rapidly converge to zero. For instance, the error bounds are only 1.8% and 9.7%, respectively, when $\Delta t \simeq 0.2/|\lambda|_{max}$ which is the five times finer sampling interval. The application of Eq.(2.7) to the CARMA model is equivalent to applying Euler formula to the controllable canonical form of its state space model, since Eq.(2.7) is just a stack of Euler formula upto the higher order derivatives. Therefore, Lemma 9 ensures the consistency of the estimator approximated by Eq.(2.7). The conditions (b) and (c) in Assumption 4 are from the well-known Nyquist theorem (Soderstrom, 1991; Mahata & Fu, 2007). The condition (c) has already been used in Lemma 4. Though we do not usually know the exact eigenvalues of A_c and the component frequency in $W(t)$ before the modeling, the conditions in Assumption 4 are useful to choose an appropriate Δt based on our background knowledge of the objective system in many practical problems.

Because Assumption 3[3] and Assumption 4 (b), (c) enables the DARMA modeling that is unbiased and consistent under a large data set (Brockwell & Davis, 1991), our approach also provides the CARMA and the SARMA modeling that is unbiased and consistent under the large data set and a small sampling interval meeting with Assumption 4 (a).

3.4 Proposed principle and algorithm

Sections 3.4.1, 3.4.2 and 3.4.3 address the objectives (3.1) and (3.2).We clarify the generic mathematical relationships between the matrices of the CARMA, the SARMA and the ARMA models, under Assumptions 3 and 4. The relationships provide a theoretical basis for their modeling. A new modeling approach for continuous time, multivariate, linear Markov systems is proposed in Section 3.4.4, which satisfies the objective (3.3).

[3]More strictly speaking, ARMA modeling requires stability of its objective system, and the state space model of minimum dimension corresponding to the ARMA model is necessarily controllable and observable (Brockwell & Davis, 1991).

3.4.1 Relationship of SARMA and DARMA models with their CARMA model

We apply the backward higher order finite difference, Eq.(2.7), to every time derivative terms, $Y^{(m)}(t)$ and $W^{(m)}(t)$, in the CARMA model, Eq.(3.3), under Assumptions 3 and 4. Then, we obtain the following Lemma 10.

Lemma 10. *Under Assumptions 2, 3 and 4, a discrete time approximation of a CARMA model is represented as follows, where $S_p = -I$ and $R_0 = I$.*

$$\begin{aligned} Y(t) &= -\left(\sum_{m=0}^{p} S_m \Delta t^{-m}\right)^{-1} \sum_{j=1}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t - j\Delta t) \\ &\quad - \left(\sum_{m=0}^{p} S_m \Delta t^{-m}\right)^{-1} \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right) W(t) \\ &\quad - \left(\sum_{m=0}^{p} S_m \Delta t^{-m}\right)^{-1} \sum_{j=1}^{q} (-1)^j \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m} W(t - j\Delta t), \end{aligned} \tag{3.10}$$

and

$$\begin{aligned} Y(t) &= \left[I + \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{m=0}^{p} S_m \Delta t^{-m}\right] Y(t) \\ &\quad + \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{j=1}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t - j\Delta t) \\ &\quad + W(t) \\ &\quad + \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{j=1}^{q} (-1)^j \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m} W(t - j\Delta t). \end{aligned} \tag{3.11}$$

■

Proof. *The proof is presented in 3.9.3.* □

We can see that Eq.(3.11) corresponds to the formula of the SARMA model in Eq.(3.2). The $p + q + 1$ coefficient matrices in the SARMA model are uniquely constrained by the $p + q$ matrices, S_m, $m = 0, \ldots, p-1$ and R_m, $m = 1, \ldots, q$, in the original CARMA model, Eq.(3.3). By comparing Eq.(3.2) and Eq.(3.11), we obtain the following representation of the coefficient matrices of a SARMA model

by the coefficient matrices of a CARMA model.

$$\Psi_0 = I + \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{m=0}^{p} S_m \Delta t^{-m}, \tag{3.12}$$

$$\Psi_j = (-1)^j \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m}, \tag{3.13}$$

$$\Omega_j = (-1)^j \left(\sum_{m=0}^{q} R_m \Delta t^{-m}\right)^{-1} \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m}, \tag{3.14}$$

where $S_p = -I$ and $R_0 = I$. For a given CARMA model, we can uniquely obtain a SARMA model by these equations. Moreover, Eq.(3.10) represents the formula of the DARMA model, Eq.(3.1), and the coefficient matrices of the DARMA model include some information of S_m and R_m from the original CARMA model.

3.4.2 Derivation of the CARMA model from the SARMA model

In the opposite way to Eq.(3.12), (3.13) and (3.14), we can now present the mathematical relationship that reproduces the CARMA model from its SARMA model under the selection of P to share the external noise term $W(t)$ between these two models as discussed in Section 3.3.1. The subsequent sections show that this constraint uniquely determines Ψ_0 of the SARMA model.

Theorem 3. *Under Assumptions 2, 3 and 4, the coefficient matrices of the CARMA model in Eq.(3.3) are represented by the coefficient matrices of the SARMA model in Eq.(3.2) as follows, where $S_p = -I$ and $R_0 = I$.*

$$\begin{aligned} S_0 &= \Delta t^{-p} I - \\ &\sum_{m=1}^{p-1} \left\{ (-1)^m \left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right)^{-1} \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j \right. \\ &\left. + (-1)^{p+m-1} \Delta t^{-p} \frac{p!}{(p-m)!m!} I \right\} + (-1)^p \Delta t^{-p} \Psi_p^{-1} (I - \Psi_0). \end{aligned} \tag{3.15}$$

$$S_m = (-1)^m \Delta t^{m-p} \left\{ \Delta t^p \left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right)^{-1} \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + (-1)^{p-1} \frac{p!}{(p-m)!m!} I \right\}, \tag{3.16}$$

where $1 \leq m \leq p-1$.

$$R_m = (-1)^m \Delta t^m \left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right)^{-1} \sum_{j=m}^{q} \frac{j!}{(j-m)!m!} \Omega_j, \tag{3.17}$$

where $1 \leq m \leq q$. ■

Proof. *The proof is presented in 3.9.4.* □

Eq.(3.12)–(3.17) indicate a bijective relationship between the CARMA and the SARMA models as discussed in Section 3.3.1.

3.4.3 Derivation of the SARMA model from the DARMA model

Next, we show the bijective correspondence of the DARMA model with system dynamics under Assumptions 3 and 4. In the meantime, we provide a mathematical relationship that constructs the SARMA model from its DARMA model.

Using Theorem 3, we can derive the following, which constructs Ψ_0 from the DARMA matrices.

Theorem 4. *Under Assumptions 2, 3 and 4, the matrix* Ψ_0 *is uniquely represented by DARMA matrices as follows.*

$$\Psi_0 = I + (-1)^{p+1} \Delta t^{-p} \Phi_p^{-1} \left(\sum_{j=1}^{q} \Theta_j \sum_{k=1}^{j} (-1)^k \frac{j!}{(j-k)!k!} - I \right). \tag{3.18}$$

■

Proof. *The proof is presented in 3.9.5.* □

Because $I - \Psi_0$ is always regular by Lemma 6, 7 and Eq.(3.12), a unique SARMA model is identified from the DARMA model through Eq.(3.7b), (3.8b) and (3.9b)

under a given Ψ_0. This fact, Theorem 4 and Lemma 5 indicate a bijective correspondence between the SARMA model and its DARMA model. In conjunction with Theorem 3, we can now immediately provide the following corollary, which establishes a novel approach to the CARMA and the SARMA modeling.

Corollary 2. *A DARMA model bijectively represents a stable, continuous time, multivariate, linear Markov system that is controllable and observable across continuous and discrete time domains under the finite difference approximation and its appropriately fine sampling interval, and has bijective correspondence with its SARMA and the CARMA models.* ■

Here, we summarize all assumptions required by our proposed method. Our study in this chapter assumes that the objective system has the characteristics of

(1) linear, (2) continuous, (3) Markov, (4) stable, (5) controllable, (6) observable, and

(7) that the approximation error of the backward higher order finite difference for the time discretization of the CARMA model is sufficiently small.

Under these assumptions, the consequences provided above enable us to obtain the SARMA model from the given DARMA model and to derive the CARMA model from the SARMA model. Similarly to the DVAR modeling in the Chapter 2, these assumptions hold in the most of scientific and engineering systems in practice and are used in the conventional DARMA modeling. This is a main reason for the wide application of the DARMA modeling. Therefore, our consequences for the ARMA modeling are considered to have a wide applicability similarly to the conventional DVAR modeling. This is a significant advantage of our framework in comparison with the past SARMA modeling methods having strongly limited applicability as explained in subsection 3.2.1.

3.4.4 Proposed algorithm

This subsection addresses our objective (3.3). Similar to the SVAR and CVAR models derivations discussed in chapter 2, we can obtain the SARMA model from a given DARMA model, and further derive the CARMA model from the SARMA

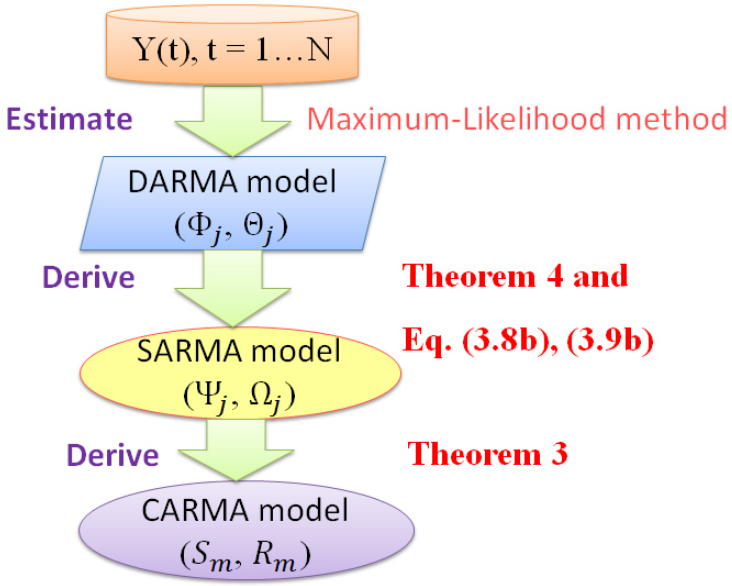

Figure 3.1: The scheme of CSARMA modeling algorithm.

model under the assumptions on the type of the objective system, Assumptions 2, 3 and 4. We can use the following algorithm to obtain these two models after we have derived the DARMA model from a given time series data set. We call the new algorithm the *continuous time structural autoregressive and moving average* (CSARMA) modeling algorithm. To estimate the DARMA model required for this algorithm, we can apply traditional techniques such as the maximum likelihood method which computes an exact likelihood function using the Kalman filter (Shea, 1987) and applies a quasi-Newton algorithm to search for the maximum of the log-likelihood solution (Gill & Murray, 1972).

Input: A DARMA model of a system, as in Eq.(3.1).

Output: SARMA, CARMA models of the system, as in Eq.(3.2) and (3.3).

1. Obtain a matrix Ψ_0 by substituting the DARMA matrices Φ_p and $\Theta_j, j = 1, \ldots, q$ into Eq.(3.18) of Theorem 4. Estimate the SARMA matrices $\Psi_j, j = 1, \ldots, p$ and $\Omega_j, j = 1, \ldots, q$ by using Eq.(3.7b), (3.8b), the given matrices $\Phi_j, j = 1, \ldots, p$, $\Theta_j, j = 1, \ldots, q$, and Ψ_0.

2. Obtain the CARMA matrices $S_m, m = 0, \ldots, p-1$ and $R_m, m = 1, \ldots, q$ by using Eq.(3.15), (3.16) and (3.17) in Theorem 3, and $\Psi_j, j = 0, \ldots, p$ and $\Omega_j, j = 1, \ldots, q$ estimated in the former step.

3.5 Performance evaluation by using artificial data

To address the objective (3.4), we have assessed the performance of our proposed CSARMA method using numerical experiments with artificial data. The first purpose of this section is to evaluate if the DARMA, the SARMA and the CARMA models derived from our artificial data match the original. Since these models have a bijective correspondence with the system dynamics, they should match the original models. The second purpose of this section is to evaluate if our approach has wider applicability and higher accuracy than a representative method that derives the SARMA model.

3.5.1 Illustrative example of CSARMA application

For this example we use the same coupled oscillator system as depicted in Fig. 2.2 and described in Section 2.4.1. For the illustration of the CSARMA method application we additionally introduce to this system the MA components by using a photo camera to observe the two deviations x_1 and x_2 in our problem. The camera has an exposure time Δt. This temporal ambiguity creates observation errors of x_1 and x_2, since the objects move over the exposure time. As the result, the camera observes their average deviations during $[t, t + \Delta t]$. Accordingly, their errors from the exact deviations is $\frac{\Delta t}{2} v_1$ for the object 1 and $\frac{\Delta t}{2} v_2$ for the object 2.

This system is exactly represented by a controllable canonical form of the state space model in continuous time domain which explicitly indicates kinematics, air friction and observation errors. It consists of the linear differential system equations having external process noises and observation equations of the state variables (Hinrichsen & Pritchard, 2005). Each of the linear differential system equations

represents an individual mechanism independently disturbed by the external noise.

$$\frac{dX}{dt} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{-2\kappa}{M} & \frac{\kappa}{M} & -\frac{c}{M} & 0 \\ \frac{\kappa}{M} & \frac{-2\kappa}{M} & 0 & -\frac{c}{M} \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} W \text{ and}$$

$$Y = \begin{bmatrix} 1 & 0 & \frac{\Delta t}{2} & 0 \\ 0 & 1 & 0 & \frac{\Delta t}{2} \end{bmatrix}^T X,$$

where $X = [x_1, x_2, v_1, v_2]^T$ and $W = [w_1, w_2]^T$. w_1 and w_1 are the external process noises of x_1 and x_2. The controllable canonical form of the state space model has direct one to one correspondence to the CARMA model (Hinrichsen & Pritchard, 2005; Brockwell & Davis, 1991), and in the case of the coupled oscillator, we have the following CARMA (2,1) model.

$$\begin{aligned} Y^{(2)}(t) = & \begin{bmatrix} \frac{-2\kappa}{M} & \frac{\kappa}{M} \\ \frac{\kappa}{M} & \frac{-2\kappa}{M} \end{bmatrix} Y^{(0)}(t) + \begin{bmatrix} -\frac{c}{M} & 0 \\ 0 & -\frac{c}{M} \end{bmatrix} Y^{(1)}(t) \\ & + W^0(t) + \begin{bmatrix} \frac{\Delta t}{2} & 0 \\ 0 & \frac{\Delta t}{2} \end{bmatrix} W^{(1)}(t). \end{aligned} \tag{3.19}$$

We gave the values of the spring constant κ=0.38 N/m, the mass M=1 kg, the air resistance coefficient for the mass $c = 0.5Ns/m$ and the exposure time $\Delta t = 1$ s, the period of oscillation $T = 2\pi\sqrt{\frac{M}{\kappa}} = 10.2$ s. Then we transformed this CARMA model to its corresponding SARMA and DARMA models by using Eq.(3.12)-(3.14) and relations in Eq.(3.7a) and (3.8a) under a time granularity $\delta t = 0.1$ s far smaller than Δt to simulate an approximately continuous process. We confirmed that the stability, the controllability and the observability of the DARMA model in this discrete time domain with δt hold to ensure the conditions required for the DARMA modeling. We further generated a time series $Y(t)$ by using the derived DARMA model parameters and Eq.(3.1). The length of the generated time series is 1000 data points. The external bivariate noises $U(t)$ were generated by using an i.i.d. $N(0, \sigma^2)$ distribution where σ is randomly chosen from [0.3,0.7] to maintain the identifiability of the DARMA model. Since the sampling time of the system is Δt

= 1 s, we further sampled the generated time series by choosing every 10th point in it.

We applied the CSARMA algorithm shown in Fig. 3.5.2 to this sampled time series and estimated the DARMA , SARMA and CARMA models. The correct models orders, p=2 and q=1, are provided for the estimation. Table A.2 shows the comparison between the original models derived from the CARMA model under the discretization with $\Delta t = 1$ s and the estimated models. The table indicates that the SARMA and CARMA models estimated by the CSARMA approach match well with their original models. We also see that the original DARMA model matches well with the DARMA model estimated by Maximum-Likelihood method. The last corresponds to our conclusion in the previous section that the objective system is represented by a unique DARMA model when it is linear Markov, stable, controllable and observable system in continuous time domain. Accordingly, the CSARMA modeling appropriately reconstructs the original SARMA and CARMA models of the system from a given time series, and provides the correct canonical relation between the variables in the original system.

3.5.2 Accuracy of the proposed method

First, we performed a set of computer simulations to evaluate the accuracy of the DARMA, the SARMA and the CARMA models derived using our CSARMA method. The procedure of the numerical experiments is shown in Fig. 3.2.

In Block (1), we artificially generated the matrices of a CARMA model, $S_m, j = 0, \ldots, p-1$, $R_m, j = 1, \ldots, q$ $(q < p)$. The elements of the matrices were generated by a uniformly distributed random value in the interval $(-1.5, 1.5)$. Then, the stability, the controllability and the observability of the generated model were checked in the continuous time domain for Assumption 3. The conditions of Assumption 4 are also checked for a sampling interval $\Delta t = 1$. We repeated the generation of the matrices until we obtained a CARMA model satisfying all of

Table 3.1: The parameter matrices of the original and estimated DARMA , SARMA and CARMA models, where Z and $\hat{Z}$ represent original and estimated matrices, respectively, for $Z = \Phi, \Theta, \Psi, \Omega, S$ and R.

DARMA model			SARMA model			CARMA model		
Φ_1	1.138	0.191	Ψ_0	−0.507	0.253	S_0	−0.760	0.380
	0.191	1.138		0.253	−0.507		0.380	−0.760
$\hat{\Phi}_1$	1.137	0.193	$\hat{\Psi}_0$	−0.494	0.252	$\hat{S}_0$	−0.763	0.384
	0.194	1.139		0.260	−0.512		0.388	−0.758
Φ_2	−0.455	−0.076	Ψ_1	1.67	0	S_1	−0.5	0
	−0.076	−0.455		0	1.67		0	−0.5
$\hat{\Phi}_2$	−0.453	−0.076	$\hat{\Psi}_1$	1.65	0	$\hat{S}_1$	−0.51	0
	−0.077	−0.453		0	1.68		0	−0.51
Θ_1	−0.333	0	Ψ_2	−0.67	0	R_1	0.500	0
	0	−0.333		0	−0.67		0	0.500
$\hat{\Theta}_1$	−0.361	0	$\hat{\Psi}_2$	−0.66	0	$\hat{R}_1$	0.521	0
	0	−0.316		0	−0.67		0	0.489
			Ω_1	−0.333	0			
				0	−0.333			
			$\hat{\Omega}_1$	−0.342	0			
				0	−0.328			

these conditions. In Block (2), we computed the matrices of the SARMA model, $\Psi_j, j = 0, \ldots, p$ and $\Omega_j, j = 1, \ldots, q$, from the CARMA model using Eq.(3.12), (3.13), (3.14) and $\Delta t = 1$. In Block (3), we computed the matrices of the DARMA model, $\Phi_j, j = 1, \ldots, p$ and $\Theta_j, j = 1, \ldots, q$, from the matrices of the SARMA model using Eq.(3.7a) and Eq.(3.8a). Then, we further checked the stability, the controllability and the observability of the DARMA model in the discrete time domain, and retained the model only if all these conditions were met. By applying the double check of the three conditions, we tested if the time discretization using time granularity $\Delta t = 1$ preserved the conditions required for the ARMA model-

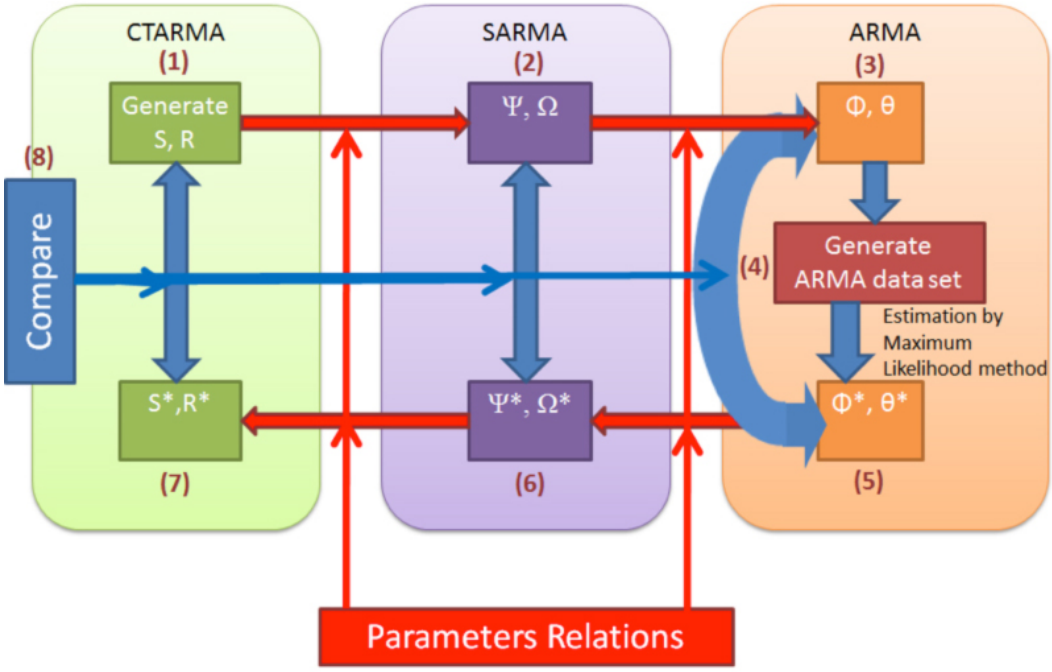

Figure 3.2: Procedure of our experiment.

ing. If they were not preserved, we repeated the entire model generation process from Block (1) to Block (3), until we obtained the CARMA, the SARMA and the DARMA models satisfying the conditions. This generation procedure qualitatively ensures Assumption 4. We checked the three conditions using the state space model explained in Section 3.3.1. The model is stable if all the eigenvalues of the matricies A_c (and A_d) have negative real parts (lie inside the unit circle on a complex number plain) (Brockwell & Davis, 1991; Stamer et al., 1996). We checked the controllability and the observability conditions by confirming if $[B_c \ A_c B_c \dots A_c^{dp-1} B_c]$ ($[B_d \ A_d B_d \dots A_d^{dp-1} B_d]$) and $[C_c \ C_c A_c \dots C_c A_c^{dp-1}]^T$ ($[C_d \ C_d A_d \dots C_d A_d^{dp-1}]^T$) are row and column full rank, respectively (Brockwell & Davis, 1991).

In Block (4), we generated the multivariate DARMA time series data, Y(t), by using the $\Phi_j, j = 1, \dots, p$ and $\Theta_j, j = 1, \dots, q$ provided in block (3), a multivariate i.i.d. Gaussian time series U(t), and Eq.(3.1). The mean value of each element in U(t) was set to zero, and its standard deviation was randomly chosen from $[0.3, 0.7]$

to maintain the identifiability of the DARMA model.

In Block (5), we estimated the DARMA matrices, $\Phi_j, j = 1, \ldots, p$, $\Theta_j, j = 1, \ldots, q$ as $\hat{\Phi}_j, j = 1, \ldots, p$ and $\hat{\Theta}_j, j = 1, \ldots, q$ from the generated multivariate time series, using a representative multivariate DARMA modeling algorithm. We used the maximum likelihood method (Shea, 1987; Gill & Murray, 1972) implemented in a MATLAB tool named NAG (NAG, 2013) to derive the ARMA matrices. In Block (6), we estimated the SARMA matrices, $\Psi_j, j = 0, \ldots, p$ and $\Omega_j, j = 1, \ldots, q$ as $\hat{\Psi}_j, j = 0, \ldots, p$ and $\hat{\Omega}_j, j = 1, \ldots, q$, by following Step 1 in the CSARMA algorithm. Subsequently, in Block (7) we estimated the CARMA matrices, $S_m, j = 0, \ldots, p-1$ and $R_m, j = 1, \ldots, q$ as $\hat{S}_m, j = 0, \ldots, p-1$ and $\hat{R}_m, j = 1, \ldots, q$ in Step 2 of CSARMA.

In Block (8), we evaluated the accuracy of the estimated matrices compared to the original matrices. Here, we used the cosine measure of a matrix estimation accuracy presented in Eq.(2.18) to evaluate the accuracy of $X_k = \Phi_j$, Θ_j, Ψ_j, Ω_j, S_m or R_m, where $k = j$ or m.

We chose $d = 5$ as the default parameter setting for the dimension of $Y(t)$, $N = 1000$ as the number of time steps of $Y(t)$, and $p = 2$ and $q = 1$ as the orders of the CARMA model. Then, we assessed the estimation accuracy over various values of each parameter, while setting the other parameters to their default values. For every parameter setting, we repeated 20 experiments and evaluated the 20 accuracies, A_X, for each experiment.

Figures 3.3, 3.4, 3.5 and 3.6 show comparisons of the estimation accuracy over the various values of every parameter. In Fig. 3.3(a), each line shows the accuracies of the AR matrix estimations from each model, averaged over the 20 experiments, for multiple dimensions of Y(t), $d = 3, 5$ and 7. The error bars represent the standard deviations. Figure 3.3(b) shows similar results for the MA matrices. We observe that the accuracies of the DARMA matrices obtained using the traditional maximum likelihood method (A_Φ and A_Θ), and the ones of the SARMA and the CARMA matrices estimated using the CSARMA method (A_Ψ, A_Ω, A_S and A_R), are close to one in all cases. That is, the accuracies are high for all dimensions of the observation vectors. In addition, the accuracy slightly degrades when the dimension d is large. This is because the number of elements to be estimated in

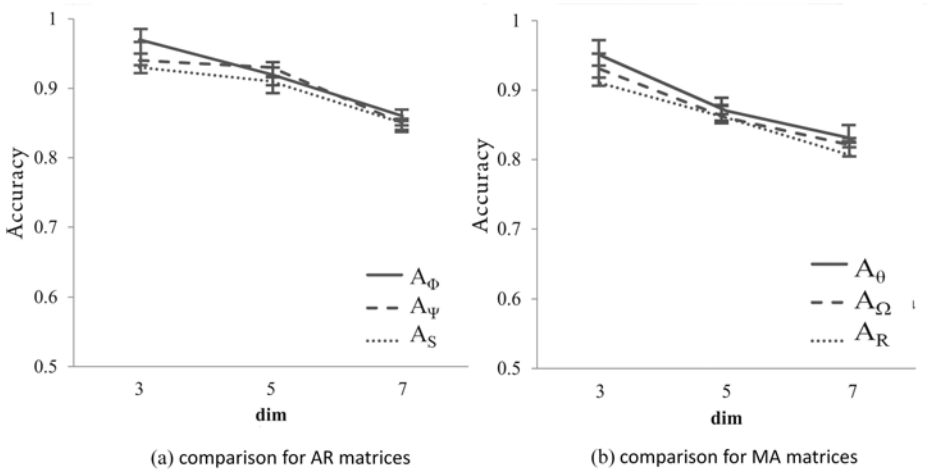

Figure 3.3: Accuracy over different dimensions d, when $N = 1000$, $p = 2$ and $q = 1$.

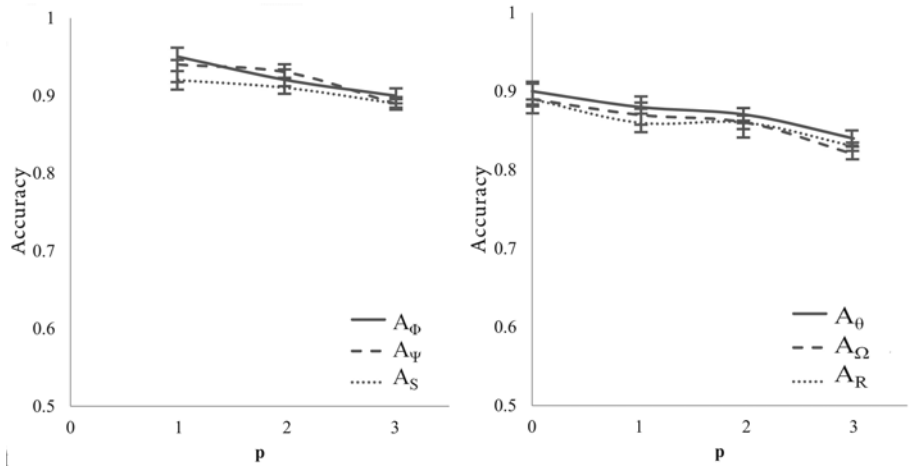

Figure 3.4: Accuracy over different AR orders p, when $d = 5$, $N = 1000$ and $q = 1$.

the AR and MA matrices is $O(d^2)$, which makes the estimation of the DARMA model statistically unstable under the same length of the given time series data, $N = 1000$.

Figure 3.4 shows the results for different AR orders, $p = 0, 1, 2, 3$, with the MA order $q = 1$. Figure 3.5 shows the results for different MA orders, $q = 0, 1, 2, 3$, with the AR order $p = 4$, where p is increased from its default value to maintain $q < p$. The accuracies of all three estimated models are high in all cases. The model estimation of the higher orders of p and q shows some degraded accuracies, as the model becomes more complicated and more elements need to be estimated. We also notice that the estimation of the MA matrices is less accurate than the

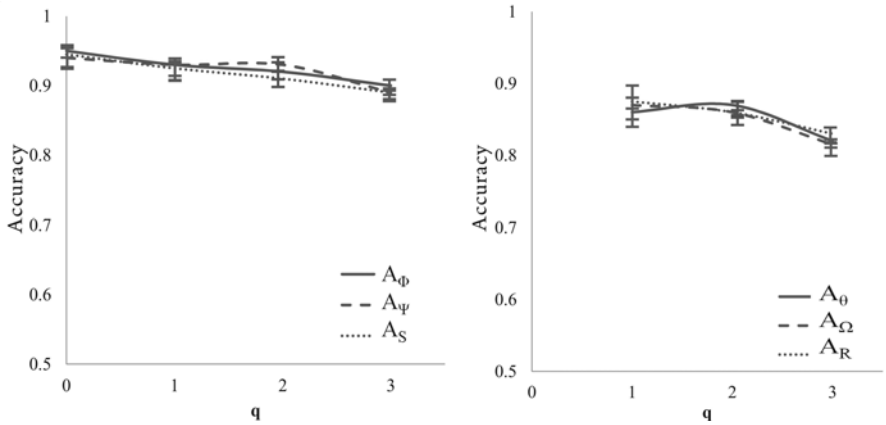

Figure 3.5: Accuracy over different MA orders q, when $d = 5$, $N = 1000$ and $p = 4$.

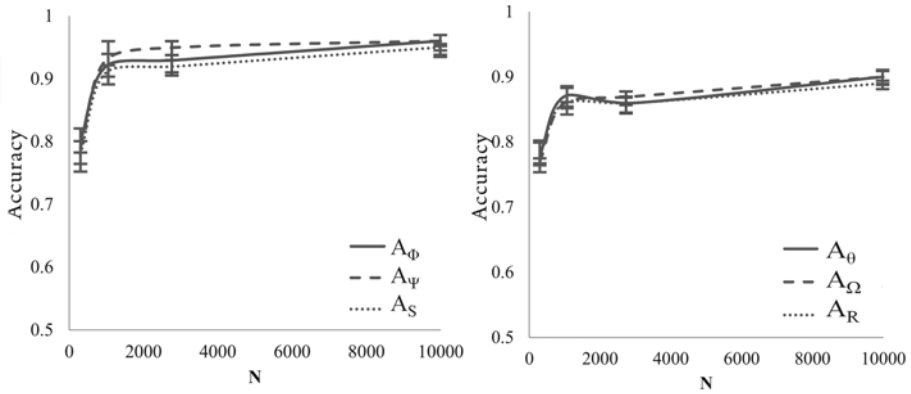

Figure 3.6: Accuracy over different steps N, when $d = 5$, $p = 2$ and $q = 1$.

AR matrices. A cause of this accuracy degradation of the MA part is considered to be less consistency of the second term associated with $W(t)$ in Lemma 9. Its other cause may be that, unlike the variables of the AR part, the noise in the MA part is not observed. The noise and the MA matrices need to be estimated at the same time, which add complications and statistical instabilities.

Figure 3.6 shows the results for $N = 300, 1000, 3000$, and 10000. The accuracies of the three estimated ARMA matrices are high for all time steps. The accuracy for $N = 300$ is not very high because of the statistical limitations of the maximum likelihood estimation, but the accuracies grow under higher N.

In summary, our proposed CSARMA method accurately captures the system dynamics and structure in the forms of the SARMA and the CARMA models.

3.5.3 Comparison of the proposed method with ARMA-LiNGAM

We have evaluated the applicability of CSARMA by comparing it to an existing representative method for deriving the SARMA model. As explained in Section 3.2, the most approaches are applicable to VAR processes but not ARMA processes, and thus they cannot be compared with our approach. Two approaches are applicable to the ARMA process (Kawahara et al., 2011; Mainassara & Francq, 2009), but the method of the parameterized SARMA model (Mainassara & Francq, 2009) is not naturally compared with our algorithm, as it requires a lot of prior knowledge to represent the system structure by a small number of parameters. Thus, we have compared our algorithm with the ARMA-LiNGAM method (Kawahara et al., 2011). This approach requires assumptions of acyclicity and non-Gaussianity. It first derives the time series of the external noises, $U(t)$, by using a maximum likelihood estimation of the DARMA model. It then applies the ICA to $U(t)$ to derive the ordering information of the variables in $Y(t)$, together with the strictly lower-triangular matrix Ψ_0.

Similar to the Section 2.4.3, to compare the two approaches in case of ARMA processes, we prepared four sets of artificial data, one for the non-Gaussian and acyclic case, one for the Gaussian and acyclic case, one for the non-Gaussian case without the acyclicity assumption and one for the Gaussian case without the acyclicity assumption. The generation process was similar to one described in Section 2.4.3. For the non-Gaussian case, we generated data by independently drawing the noise values in $W(t)$ from Gaussian distributions and subsequently passing them through a power non-linearity (raising the absolute value to an exponent in the interval $[0.5, 0.8]$ or $[1.2, 2.0]$, but keeping the original sign) to make them non-Gaussian (Shimizu et al., 2011). For the Gaussian case, we used the same procedure as in Section 3.5.2. For the acyclic case, we first randomly generated the SARMA matrices, where the matrix Ψ_0 is strictly lower-triangular, which is the necessary condition of acyclicity. Then, we checked the regularity of the matrix $\sum_{m=0}^{q} R_m \Delta t^{-m}$ in Lemma 7 and of matrix Ψ_p in Eq.(3.15) to ensure the existence of a CARMA model corresponding to the generated SARMA model. If the CARMA

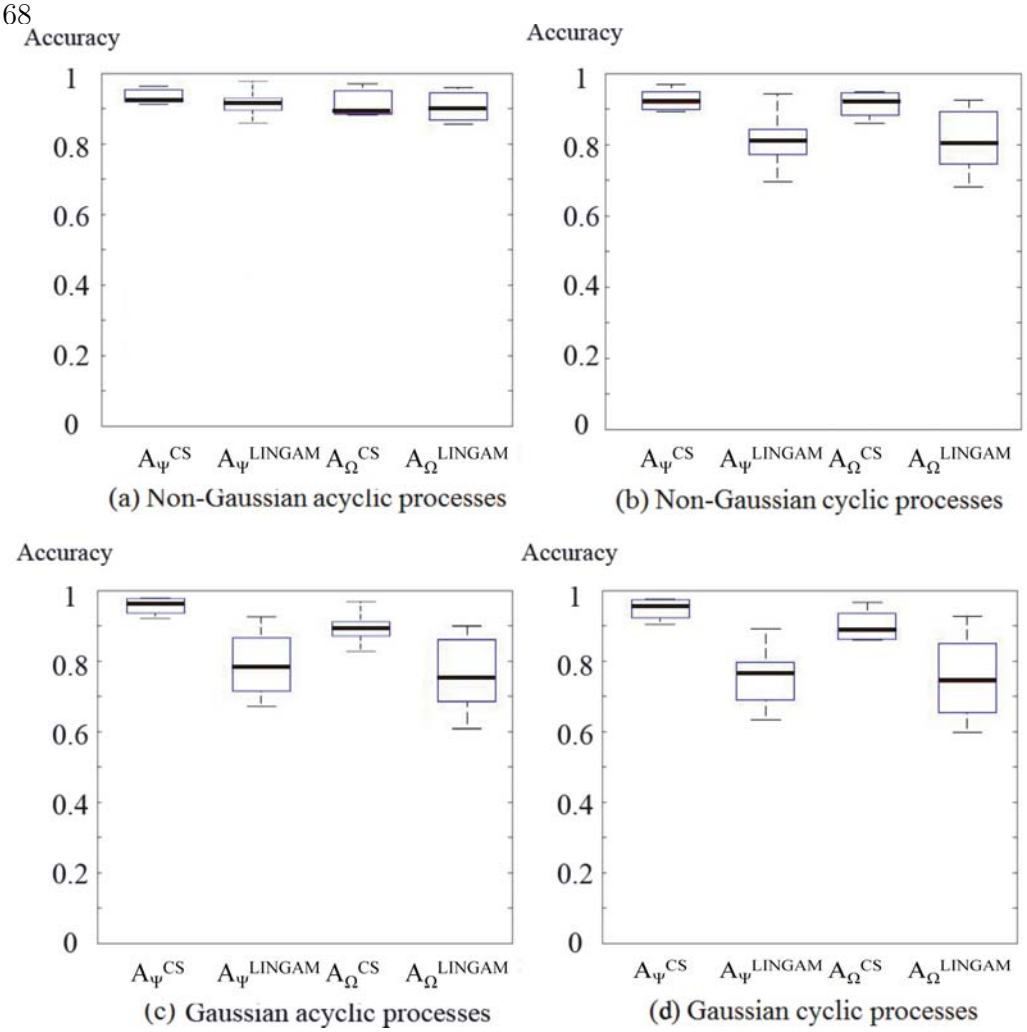

Figure 3.7: Accuracies of SARMA model estimations by CSARMA and ARMA-LiNGAM for different processes under $d = 5$, $N = 1000$, $p = 2$ and $q = 1$.

model existed, we continued the model generation process in the same way as described in Section 3.5.2. For the processes without the acyclicity assumption, we used the same generation algorithm shown in Section 3.5.2.

We applied both our proposed algorithm and the ARMA-LiNGAM algorithm to these sets and obtained different structural ARMA matrices, $\Psi_j^{CS}, j = 0, \ldots, p$, $\Omega_j^{CS}, j = 1, \ldots, q$, $\Psi_j^{LiNGAM}, j = 0, \ldots, p$, and $\Omega_j^{LiNGAM}, j = 1, \ldots, q$. Then, we compared the obtained matrices to the originally generated ones using Eq.(2.18), similarly to Section 3.5.2. The results are shown in Fig. 3.7. Figure 3.7(a) shows the box plots of the estimation accuracies of CSARMA and ARMA-LiNGAM for the

non-Gaussian and acyclic processes. Both methods are very accurate for the non-Gaussian and acyclic processes. However, when the assumptions of non-Gaussianity and acyclicity are not met as shown in Fig. 3.7(b), (c) and (d), the accuracy of the ARMA-LiNGAM method is substantially lower than the CSARMA approach. As expected, the applicability of ARMA-LiNGAM is limited to the non-Gaussian and acyclic processes. In contrast, our proposed CSARMA approach is widely applicable to the continuous time, linear Markov system as far as the system is stable, controllable, observable, and observed using an appropriately fine sampling interval.

3.6 Performance evaluation using real world data

In this section, we evaluate the practicality of our CSARMA approach to achieve the objective (3.4). We applied the CSARMA algorithm to real world experimental data. We used a reactor noise time series, measured in an impulse fast neutron research reactor called IBR-2 at the Joint Institute of Nuclear Research in Dubna, Russia, which is depicted in Fig. 2.8 and described in Section 2.4.4 (Pepyolyshev, 1988).

The aim of this analysis is to identify the influences between the peak power and the neutron reflections of the main and the additional reflectors over their sequential operations, but not to identify the dynamics to generate an individual power peak. Accordingly, our objective experimental data include time series of peak values of the power pulses, Q, axial deviations of the main neutron reflector, X_Q, and the additional neutron reflector, X_A, measured in parallel at every instant of the power pulse during the stable reactor operation. The axial deviations of the reflectors are their angular deviation from the vertical central line of the reactor core. The sampling frequency of the time series data was necessarily equal to the frequency of the pulse operation of IBR-2, which is 5Hz (Pepyolyshev, 1988), and thus $\Delta t = 0.2$ (sec). This is sufficiently fine sampling to capture the dynamics of the peak powers over their sequence, because the peak power and its associated neutron reflection change more slowly by breading, accumulation and leakage of the neutrons in the reactor core. The time series data contained 2048 time steps. Each variable was

normalized to give a zero mean and a unit standard deviation.

The dynamics of this reactor are known to be approximately represented by CARMA(2,1). The process of the heat removal from the core and the negative feedback process of core temperature to power generation are approximately represented by the second order delay process. In addition, the impacts of the reflectors are represented by the first order delay processes of the external influences. Accordingly, we estimated the DARMA(2,1) from the reactor noise time series using the maximum likelihood method, and then calculated its SARMA and CARMA matrices using our CSARMA method.

Eq.(3.20) shows the AR and MA matrices of the DARMA model estimated from the time series data. Though the DARMA model bijectively corresponds to the system dynamics of the IBR-2 reactor as noted in Corollary 2 , these matrices do not indicate clear structures reflecting the dynamics. For example, the impact of the main neutron reflector to the peak power is known to be 4 times larger than that of the additional reflector as mentioned earlier. However, the ratios of the $(1, 2)$ and $(1, 3)$-elements in Φ_1 and Φ_2 are far from 4, respectively. This is because the DARMA model is not structural, and its coefficient matrices are not decomposed into individual effects as explained in Section 3.2.1.

$$
\begin{aligned}
\Phi_1 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} -0.1265 & -0.0905 & 0.0074 \\ -0.0029 & -0.0932 & -0.0400 \\ -0.0270 & 1.1384 & 0.2169 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
\Phi_2 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 0.3283 & -0.2414 & -0.2597 \\ -0.0092 & -0.4051 & 0.0893 \\ -0.0012 & 2.4921 & -0.0659 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
\Theta_1 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} -0.4884 & -0.2947 & -0.1317 \\ 0.0013 & -0.5634 & -0.0067 \\ 0.0955 & 0.9228 & 0.4303 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}.
\end{aligned}
\tag{3.20}
$$

Eq.(3.21) shows the AR and MA matrices of the SARMA model.

$$
\begin{aligned}
\Psi_0 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 49.7402 & 150.2176 & 65.0670 \\ 1.2771 & 15.7391 & 16.4848 \\ 11.1788 & 204.5666 & 80.6094 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
\Psi_1 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 4.8445 & -55.6610 & -8.4650 \\ -0.2408 & -17.2771 & -2.9954 \\ -0.1421 & -70.5501 & -9.1673 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
\Psi_2 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} -14.5413 & -89.5345 & 3.5313 \\ -0.2639 & -34.8028 & 0.1018 \\ -1.6924 & -112.8261 & -10.1184 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array},
\end{aligned}
\tag{3.21}
$$

$$
\Omega_1 = \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} -0.4183 & 3.5814 & -0.1413 \\ 0.0106 & 0.3921 & -0.0041 \\ 0.0677 & 4.5130 & -0.5953 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}.
$$

The major effects between the variables are almost localized at $(1,2)$, $(1,3)$ and $(3,2)$-elements in every matrix. In particular, $(1,2)$ and $(1,3)$-elements in Ψ_0 indicate the strong impacts of the two neutron reflectors to the peak power which are caused by the neutron generation within the time interval Δt. The ratio of the $(1,2)$-element over the $(1,3)$-element in Ψ_0 is almost (2.3) reflecting their different effects. However, this ratio is less than the designed impact ratio of the two reflectors. As we will quantitatively discuss later, this is because of the impact amplification of the additional reflector in the time interval Δt by neutron breading process. In addition, $(1,2)$-elements in Ψ_1 and Ψ_2 are negative because of the negative feedback of the peak power. Once the neutron population is increased in the core by the neutron reflector, the core temperature is increased through the activation of the nuclear fission chain reaction. Since the high temperature has an effect to reduce the efficiency of the individual nuclear fission reaction, this suppresses the power generation. This negative feedback forms a second order time delay dynamics generating oscillatory behaviors of the variables reflected by the negative signs of the main reflectors' impacts in Ψ_1 and Ψ_2. Further, we notice that some dependency of the additional reflector's motion to the main reflector at $(3,2)$-elements in all the matrices. Though the cause of this dependency is not clear, the voltage deriving the motor of the additional reflector could be influenced by the motor operation of the main reflector sharing an electricity power supply. We observe little impact from the peak power to the reflectors' motions in all the matrices. This is consistent with the design of the IBR-2 reactor, where the rotation of the reflectors driven by motors does not depend on the reactor's power.

Finally, Eq.(3.22) shows the AR and MA matrices of the CARMA model.

$$
\begin{aligned}
S_0 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 69.3777 & 257.7026 & 53.5547 \\ 0.0604 & -27.8419 & 9.4582 \\ 10.8091 & 319.7046 & 34.6218 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
S_1 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 7.9950 & 5.5802 & 0.2084 \\ 0.0494 & 12.4589 & 0.4280 \\ -0.1450 & 6.5110 & 9.7232 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}, \\
R_1 &= \begin{array}{c} \begin{array}{ccc} Q & X_Q & X_A \end{array} \\ \left[\begin{array}{ccc} 0.1540 & -1.2698 & 0.1108 \\ -0.0028 & -0.0509 & 0.0005 \\ -0.0284 & -1.4496 & 0.2697 \end{array}\right] \end{array} \begin{array}{c} Q \\ X_Q \\ X_A \end{array}.
\end{aligned}
\tag{3.22}
$$

This CARMA model more clearly represents the processes working in the system dynamics than the SARMA model. We see that the major effects between the variables are almost localized at $(1,2)$, $(1,3)$ and $(3,2)$-elements in S_0 and S_1 similarly to the SARMA model. The $(1,2)$ and the $(1,3)$-elements in S_0 represent the impacts of the main and the additional reflector's motions to the acceleration of the peak power, respectively. Their ratio is almost 4.8 which is roughly consistent with the designed impact ratio of these two reflectors. We also observe some dependency of the additional reflector's motion to the main reflector's at the $(3,2)$-element in S_0 and S_1. This was also seen in the SARMA model. The other elements in these matrices are negligibly small. All these consequences are consistent with the dynamics of the IBR-2 reactor.

By using the closed-form expression, Eq.(3.12), relating Ψ_0 of the SARMA model with the matrices of the CARMA model, we can quantitatively analyze the difference of the impact ratios between the two reflectors across the two models. The expression for the CARMA$(2,1)$ and the SARMA $(2,1)$ models is written as

$$
\Psi_0 = (I + \Delta t^{-1} R_1)^{-1}(S_0 + \Delta t^{-1} S_1 - \Delta t^{-2} I) + I. \tag{3.23}
$$

The two factors of the first term in the r.h.s. are represented as follows by substi-

tuting Eq.(3.22).

$$(I + \Delta t^{-1} R_1)^{-1} = \begin{bmatrix} 0.5817 & 3.5814 & -0.1413 \\ 0.0106 & 1.3921 & -0.0041 \\ 0.0677 & 4.5130 & 0.4047 \end{bmatrix}, \quad (3.24)$$

$$S_0 + \Delta t^{-1} S_1 - \Delta t^{-2} I = \begin{bmatrix} 84.3529 & 285.6034 & 54.5968 \\ 0.3073 & 9.4525 & 11.5980 \\ 10.0843 & 352.2594 & 58.2379 \end{bmatrix}. \quad (3.25)$$

These matrices indicate that the $(1,2)$-element of Ψ_0 in Eq.(3.21) mainly comes from the product of $(1,1)$-element in Eq.(3.24) and the $(1,2)$-element in Eq.(3.25), where the former reflects the process between Q and its intrinsic disturbance, and the latter is the direct influence of X_Q to Q. Thus, the main reflector directly impacts the peak power. In contrast, $(1,3)$-element of Ψ_0 in Eq.(3.21) mainly consists of the product of $(1,1)$-element in Eq.(3.24) and the $(1,3)$-element in Eq.(3.25) and the product of $(1,2)$-element in Eq.(3.24) and the $(2,3)$-element in Eq.(3.25). The first product represents the direct influence of X_A to Q similarly to the case of the main reflector, while the second product shows that X_A influences Q through some process associated with X_Q. This reflects the mechanisms of neutron breading in the reactor. The angle deviations of the two reflectors are independent in principle. However, the neutron population is exponentially increased during the finite time interval Δt by the impacts of the neutron reflectors through the nuclear chain reaction process. In particular, the breading becomes highly active, when the impact of the additional reflector is added to that of the main reflector. This process amplify the effect of the additional reflector under the strong effect of the main reflector in Δt. The second product reflects this process, and this is the reason that the impact ratio between the two reflectors in the SARMA model is smaller than that in the CARMA model.

In summary, the application of our CSARMA approach gives us rich information about the physical structure governing the variables. In particular, it provides explicit information on the relation between the fundamental system mechanisms represented by time differential equations and the structural dynamics appearing in the time series sampled at discrete time steps. This is possible through the

mathematical construction of the SARMA and the CARMA models out of the DARMA model using the closed-form expressions.

3.7 Discussion

Many techniques have been used to study the SARMA modeling of multivariate, linear Markov systems, as we reviewed in Section 3.2. However, very few studies in statistical causal inference have investigated the structural and/or causal modeling of dynamic Markov systems from given time series data. Research combining AR/ARMA modeling with the LiNGAM approach spans both fields (Hyvarinen et al., 2008; Kawahara et al., 2011). In (Voortman et al., 2010), the authors proposed a method to learn the causal structures of continuous time Markov systems in the framework of statistical causal inference. Though its basic framework and objective is different from our proposed method, it shares some similar features. For example, its model consists of higher order time difference variables and requires faithfulness of the objective system, ensured by excluding its equilibrium states. This feature seems to be associated with the requirement of the controllability and the observability in our approach. That is, some parts of the system can stay at the equilibrium, or their change can be unidentified, if the system is not fully controllable by the external random noises and not fully observable through the sampled time series data. Associated with these requirements, some past studies argued that the structural dependency among variables in an equilibrium model of a system is asymptotically deduced from the continuous time model of the system dynamics (Fisher, 1970; Iwasaki & Simon, 1994; Pearl, 2000; Lacerda et al., 2008; Mooij et al., 2013). Though the strict relation between this argument and the above requirements is unclear, these seem keys to connect the structures in the static and the dynamic modes of a system.

The CARMA and the SARMA models provided by this approach consist of linear time differential and difference equations of variables observed from an objective system, and are ensured to bijectively correspond to the objective system dynamics. Hence, our CSARMA approach enables us to empirically uncover scientific models and their associated laws in the system. In engineering oriented fields, our

CSARMA approach can be used to monitor and diagnose anomalous changes of an objective system through time series measurements. For example, in the case of the IBR-2 reactor presented in Section 3.6, the integrity of the pulse power operation in the reactor can be monitored by periodic applications of our approach. If we observe some anomalous changes at $(1, 2)$ and $(1, 3)$ elements of the CARMA and the SARMA matrices, we know about the onset of some defects of the neutron reflectors and/or the neutron generation process.

3.8 Conclusion

In this chapter, we have succeeded in achieving its four objectives. First, we showed that a DARMA model of a continuous time, multivariate, linear Markov system has a bijective correspondence with system dynamics under the finite difference approximation, if the system is stable, controllable, observable and observed using an appropriately fine sampling interval. Second, we clarified the closed-form expressions bijectively relating the CARMA, the SARMA and the DARMA models of the system under the finite difference approximation. Third, we proposed a new modeling approach to derive the CARMA and the SARMA models from the DARMA model, which is estimated using time series data observed from a system. Finally, we demonstrated the practical performance of our proposed approach through some numerical experiments using both artificial and real world data.

An issue remained is to apply a more accurate finite difference schemes than the backward scheme, Eq.(2.7). For example, the central higher order finite difference has its approximation error $O(\Delta t^2)$ that makes our models more consistent than the error $O(\Delta t)$ of Eq.(2.7) (Levy & Lessman, 1992; LeVeque, 2007). The derivation of the closed-form expression using this central scheme to bijectively relate the three ARMA models is our future work. In addition, more rigorous relationships of the ARMA models should be assessed by using the state space representations of the ARMA process across the continuous and the discrete time domains. Particularly, the existence of the bijective relationships between the CARMA(p, q) and the DARMA$(p, p-1)$ models is an important issue to be explored in the future.

Nonetheless, the developed CSARMA modeling approach is a strong instrument

for the analysis of the real world systems. In the next chapter we show its application for the analysis of the complex physical system, such as a Boiling Nuclear Rector.

3.9 Appendix

3.9.1 The proof of Lemma 7

Proof. Because of the controllability of the system, the CARMA model of Eq.(3.3) has the controllable canonical form of its state space model, Eq.(3.5). As the system is also observable, we can rewrite the controllable canonical form into its observable canonical form,

$$\begin{aligned} dX_c'(t)/dt &= A_c^T X_c'(t) + C_c W(t), \\ Y(t) &= B_c^T X_c'(t). \end{aligned}$$

The state variable vector $X_c'(t)$ in the observable canonical form and $X_c(t)$ in the controllable canonical form have the relation $X_c'(t) = TX_c(t)$, where T is a regular matrix. Thus, B_c and C_c have the relation $C_c = TB_c$. Because B_c is full column rank by its definition, so is C_c.

Now, take a d-dimensional vector $W_i^{(m)}(t) = [0 \ \ldots \ 0 \ w_i^{(m)}(t) \ 0 \ \ldots \ 0]^T$, with i-th element $w_i^{(m)}(t) = \Delta t^{-m} exp(t/\Delta t)$, and zeros elsewhere. Then, we define a dp-dimensional vector $E_i(t) = [W_i^{(0)}(t)^T \ldots W_i^{(p-1)}(t)^T]^T$. Since the vectors $E_i(t), i = 1, \ldots, d$ are mutually independent for any d and p by the definition of $W_i^{(m)}(t)$, the $dp \times d$ matrix $[E_1(t) \ldots E_d(t)]$ is full column rank. As both C_c and $[E_1(t) \ldots E_d(t)]$ are full column rank, using the definition of C_c, the $d \times d$ matrix $C_c^T[E_1(t) \ldots E_d(t)]$ $= \sum_{m=0}^{q} R_m \Delta t^{-m} \ exp(t/\Delta t)$ is full rank, *i.e.*, regular, for any t. This implies that $\sum_{m=0}^{q} R_m \Delta t^{-m}$ where $t = 0$ is also regular. □

3.9.2 The proof of Lemma 9

Proof. By applying Euler formula to the time derivatives in the first equation of Eq.(3.5), we obtain the following formula.

$$X_c(t) = (I + A_c \Delta t) X_c(t - \Delta t) + \Delta t B_c W(t).$$

We notice that this is equal to the first equation of Eq.(3.6) with the truncation of the terms more than the first order Δt. Hence, we can represent the approximation errors of the coefficient matrices in its first and the second terms, respectively, as

$$\begin{aligned} \exp(A_c\Delta t) - (I + A_c\Delta t) &= E(\exp(\Lambda\Delta t) - I - \Lambda\Delta t)E^{-1}, \\ A_c^{-1}(\exp(A_c\Delta t) - I)B_c - \Delta t B_c &= E\Lambda^{-1}(\exp(\Lambda\Delta t) - I - \Lambda\Delta t)E^{-1}B_c, \end{aligned}$$

by introducing the eigenvalue decomposition $A_c = E\Lambda E^{-1}$, where $\Lambda = diag(\lambda_1, \ldots, \lambda_d)$, and E is an eigenvector matrix. Each diagonal element of $\exp(\Lambda\Delta t) - I - \Lambda\Delta t$ is represented as

$$\exp(\lambda_j\Delta t) - 1 - \lambda_j\Delta t = \sum_{k=2}^{\infty}(\lambda_j\Delta t)^k/k!,$$

for $j = 1, \ldots, d$. Thus, it is bounded as

$$\begin{aligned} |\exp(\lambda_j\Delta t) - 1 - \lambda_j\Delta t| &\leq \sum_{k=2}^{\infty}(|\lambda_j|\Delta t)^k/k! \leq \sum_{k=2}^{\infty}(|\lambda|_{max}\Delta t)^k/k! \\ &= \exp(|\lambda|_{max}\Delta t) - 1 - |\lambda|_{max}\Delta t. \end{aligned}$$

Accordingly, the approximation error of every element in the first term is bounded by the first formulae given in this lemma. By taking into account the definition of B_c consisting of zeros and a unit submatrix, the error of every element in the second term is bounded by the second formula. □

3.9.3 The proof of Lemma 10.

Proof. By substituting Eq.(2.7) into Eq.(3.3) under Assumption 4, we obtain the following.

$$\begin{aligned} &\frac{1}{\Delta t^p}\sum_{j=0}^{p}(-1)^j\frac{p!}{(p-j)!j!}Y(t - j\Delta t) = \\ &\sum_{m=0}^{p-1}S_m\frac{1}{\Delta t^m}\sum_{j=0}^{m}(-1)^j\frac{m!}{(m-j)!j!}Y(t - j\Delta t) \\ &+W(t) + \sum_{m=1}^{q}R_m\frac{1}{\Delta t^m}\sum_{j=0}^{m}(-1)^j\frac{m!}{(m-j)!j!}W(t - j\Delta t). \end{aligned} \tag{3.26}$$

The terms of $Y(t-j\Delta t)$ are summarized with $S_p = -I$ and rewritten by permuting the summations on m and j.

$$\sum_{m=0}^{p} S_m \Delta t^{-m} \sum_{j=0}^{m} (-1)^j \frac{m!}{(m-j)!j!} Y(t-j\Delta t) =$$
$$\sum_{j=0}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t-j\Delta t).$$

Similarly, the terms of W(t) and $W(t-j\Delta t)$ are reformulated with $R_0 = I$ as follows.

$$\sum_{m=0}^{q} R_m \Delta t^{-m} \sum_{j=0}^{m} (-1)^j \frac{m!}{(m-j)!j!} W(t-j\Delta t) =$$
$$\sum_{j=0}^{q} (-1)^j \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m} W(t-j\Delta t).$$

By substituting these relationhips into Eq.(3.26), we obtain the following.

$$-\sum_{m=0}^{p} S_m \Delta t^{-m} Y(t) = \sum_{j=1}^{p} (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m} Y(t-j\Delta t)$$
$$+\sum_{m=0}^{q} R_m \Delta t^{-m} W(t) + \sum_{j=1}^{q} (-1)^j \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m} W(t-j\Delta t),$$

where $S_p = -I$ and $R_0 = I$.

By Assumption 3 and Lemma 6, $\sum_{m=0}^{p} S_m \Delta t^{-m}$ is a regular matrix. Then, by multiplying both sides of the equation by $-(\sum_{m=0}^{p} S_m \Delta t^{-m})^{-1}$, we obtain Eq.(3.10). Assumption 3 and Lemma 7 state that $\sum_{m=0}^{q} R_m \Delta t^{-m}$ is a regular matrix, so we can multiply both sides of Eq.(3.10) by $-(\sum_{m=0}^{q} R_m \Delta t^{-m})^{-1} \sum_{m=0}^{p} S_m \Delta t^{-m}$ and add Y(t) to obtain Eq.(3.11). □

3.9.4 The proof of Theorem 3.

Proof. First, we prove Eq.(3.17). Eq.(3.14) which holds under Assumptions 3 and 4 can be reformulated as

$$V\Omega_j = (-1)^j \sum_{m=j}^{q} \frac{m!}{(m-j)!j!} R_m \Delta t^{-m}, \tag{3.27}$$

where

$$V = \sum_{m=0}^{q} R_m \Delta t^{-m}, \tag{3.28}$$

which is independent of m. V is regular by Assumption 3 and Lemma 7.

In the case of j=q in Eq.(3.27), we obtain

$$V\Omega_q = (-1)^q R_q \Delta t^{-q} \Rightarrow R_q \Delta t^{-q} = (-1)^q V \Omega_q. \tag{3.29}$$

Assume the following expression for the general case, i.e., $1 \leq m \leq q$.

$$R_m \Delta t^{-m} = (-1)^m V \left(\sum_{j=m}^{q} \frac{j!}{(j-m)!m!} \Omega_j \right). \tag{3.30}$$

Note that Eq.(3.30) subsumes Eq.(3.29). When $j = k-1$, we can rewrite Eq.(3.27) as

$$\begin{aligned} V\Omega_{k-1} &= (-1)^{k-1} \sum_{m=k-1}^{q} \frac{m!}{(m-(k-1))!(k-1)!} R_m \Delta t^{-m} = \\ &(-1)^{k-1} R_{k-1} \Delta t^{-(k-1)} + (-1)^{k-1} \sum_{m=k}^{q} \frac{m!}{(m-(k-1))!(k-1)!} R_m \Delta t^{-m}. \end{aligned}$$

By substituting Eq.(3.30) for $R_m \Delta t^{-m}$ in the final term, we can derive the following.

$$\begin{aligned} &V\Omega_{k-1} = (-1)^{k-1} R_{k-1} \Delta t^{-(k-1)} \\ &+(-1)^{k-1} V \sum_{m=k}^{q} (-1)^m \frac{m!}{(m-(k-1))!(k-1)!} \sum_{j=m}^{q} \frac{j!}{(j-m)!m!} \Omega_j. \end{aligned} \tag{3.31}$$

The double summations in the last term can be rewritten as

$$\begin{aligned} &\sum_{m=k}^{q} (-1)^m \frac{m!}{(m-(k-1))!(k-1)!} \sum_{j=m}^{q} \frac{j!}{(j-m)!m!} \Omega_j = \\ &\sum_{j=k}^{q} \Omega_j \sum_{m=k}^{j} (-1)^m \frac{m!}{(m-(k-1))!(k-1)!} \frac{j!}{(j-m)!m!} I. \end{aligned}$$

Thus, we write Eq.(3.31) as follows.

$$\begin{aligned} &(-1)^{k-1} R_{k-1} \Delta t^{-(k-1)} = \\ &V\Omega_{k-1} - (-1)^{k-1} V \sum_{j=k}^{q} \Omega_j \frac{j!}{(k-1)!} \sum_{m=k}^{j} (-1)^m \frac{1}{(m-(k-1))!(j-m)!} I = \\ &V\Omega_{k-1} - V \sum_{j=k}^{q} \Omega_j \frac{j!}{(j-k+1)!(k-1)!} \left(\sum_{u=0}^{j-k+1} (-1)^u \frac{(j-k+1)!}{u!(j-u-k+1)!} I - I \right), \end{aligned}$$

where $m = u + k - 1$. According to the binomial theorem,

$$\sum_{u=0}^{j-k+1} (-1)^u \frac{(j-k+1)!}{u!(j-u-k+1)!} I = (-1+1)^{j-k+1} I = 0.$$

Thus,

$$R_{k-1}\Delta t^{-(k-1)} = (-1)^{k-1} V \left(\Omega_{k-1} + \sum_{j=k}^{q} \frac{j!}{(j-(k-1))!(k-1)!} \Omega_j \right) =$$

$$(-1)^{k-1} V \sum_{j=k-1}^{q} \frac{j!}{(j-(k-1))!(k-1)!} \Omega_j.$$

By rewriting $k-1$ by $m-1$, we know that Eq.(3.30) holds for $m-1$. By induction, Eq.(3.30) holds for all $1 \leq m \leq q$.

Furthermore, by substituting Eq.(3.30) and $R_0 = I$ into Eq.(3.28), we can obtain

$$\begin{aligned} V &= I + \sum_{k=1}^{q} (-1)^k V \left(\sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right) \Rightarrow \\ & V \left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right) = I. \end{aligned}$$

Since V is regular by Assumption 3 and Lemma 7, and I is also regular,

$$\left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right),$$

is also regular. Thus, we obtain

$$V = \left(I - \sum_{k=1}^{q} (-1)^k \sum_{j=k}^{q} \frac{j!}{(j-k)!k!} \Omega_j \right)^{-1}. \tag{3.32}$$

By substituting this into Eq.(3.30), we obtain Eq.(3.17).

Next, we prove Eq.(3.16). We can reformulate Eq.(3.13) which holds under Assumptions 3 and 4 as follows.

$$V\Psi_j = (-1)^j \sum_{m=j}^{p} \frac{m!}{(m-j)!j!} S_m \Delta t^{-m}. \tag{3.33}$$

When $j = p-1$ and $S_p = -I$, we can obtain

$$\begin{aligned} & V\Psi_{p-1} = (-1)^{p-1} S_{p-1} \Delta t^{-p+1} - (-1)^{p-1} \frac{p!}{1!(p-1)!} I \Delta t^{-p}. \\ & \therefore S_{p-1}\Delta t^{-p+1} = (-1)^{p-1} V\Psi_{p-1} + pI\Delta t^{-p}. \end{aligned} \tag{3.34}$$

Assume that Eq.(3.16) is an expression for the general case, $1 \leq m \leq p-1$. Using Eq.(3.32), we can rewrite it as

$$S_m \Delta t^{-m} = (-1)^m \left(V \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + (-1)^{p-1} \frac{p!}{(p-m)!m!} I \Delta t^{-p} \right). \quad (3.35)$$

Note that Eq.(3.35) subsumes Eq.(3.34).

We write Eq.(3.13) for the case $j = p$ as

$$\Psi_p = (-1)^{p-1} V^{-1} \Delta t^{-p}. \quad (3.36)$$

Then, we can rewrite Eq.(3.35) as follows.

$$S_m \Delta t^{-m} = (-1)^m V \sum_{j=m}^{p} \frac{j!}{(j-m)!m!} \Psi_j. \quad (3.37)$$

We can rewrite Eq.(3.33) for the case $j = k-1$ by substituting Eq.(3.37) as follows.

$$\begin{aligned} V\Psi_{k-1} &= (-1)^{k-1} S_{k-1} \Delta t^{-(k-1)} \\ &+ (-1)^{k-1} \sum_{m=k}^{p} (-1)^m \frac{m!}{(m-(k-1))!(k-1)!} V \sum_{j=m}^{p} \frac{j!}{(j-m)!m!} \Psi_j. \end{aligned}$$

By changing the order of the double summation in the last term,

$$\begin{aligned} &(-1)^{k-1} S_{k-1} \Delta t^{-(k-1)} = \\ &V\Psi_{k-1} - (-1)^{k-1} V \sum_{i=k}^{p} \Psi_j \sum_{m=k}^{j} (-1)^m \frac{m!}{(m-(k-1))!(k-1)!} \frac{j!}{(j-m)!m!} I, \end{aligned}$$

is obtained. Then we further obtain the following expression, in a similar manner to Eq.(3.31).

$$\begin{aligned} &(-1)^{k-1} S_{k-1} \Delta t^{-(k-1)} = \\ &V\Psi_{k-1} - V \sum_{j=k}^{p} \Psi_j \frac{j!}{(j-k+1)!(k-1)!} \left(\sum_{u=0}^{j-k+1} (-1)^u \frac{(j-k+1)!}{(j-u-k+1)!u!} I - I \right), \end{aligned}$$

where $m = u + k - 1$. Using the binomial theorem and replacing $k-1$ with $m-1$, we can show that Eq.(3.37) holds for $m-1$. By induction, Eq.(3.37) holds for $1 \leq m \leq p-1$. Furthermore, by substituting Eq.(3.36) into Eq.(3.37), we

can obtain Eq.(3.35), and by further substituting Eq.(3.32) into it, we can obtain Eq.(3.16) for $1 \leq m \leq p-1$.

To obtain Eq.(3.15), we substitute Eq.(3.28), (3.36) and $S_p = -I$ into Eq.(3.12) under Assumptions 3 and 4.

$$I - \Psi_0 = (-1)^{p-1}\Psi_p \left(I - \sum_{m=1}^{p-1} \Delta t^{p-m} S_m - \Delta t^p S_0 \right). \tag{3.38}$$

From Eq.(3.36), Assumption 3 and Lemma 7, Ψ_p is regular. Then, we reformulate Eq.(3.38) as follows.

$$I - \sum_{m=1}^{p-1} \Delta t^{p-m} S_m - \Delta t^p S_0 = (-1)^{p-1}\Psi_p^{-1}(I - \Psi_0).$$

Thus,

$$\Delta t^p S_0 = I - \sum_{m=1}^{p-1} \Delta t^{p-m} S_m + (-1)^p \Psi_p^{-1}(I - \Psi_0).$$

Using Eq.(3.35), we can derive the following.

$$\begin{aligned} &\Delta t^p S_0 = I + (-1)^p \Psi_p^{-1}(I - \Psi_0) \\ &- \sum_{m=1}^{p-1} \left((-1)^m \Delta t^p V \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + (-1)^{p+m-1} \frac{p!}{(p-m)!m!} I \right). \end{aligned} \tag{3.39}$$

By substituting Eq.(3.32) into Eq.(3.39), we get Eq.(3.15). □

3.9.5 The proof of Theorem 4.

Proof. Consider $\sum_{m=1}^{p-1} S_m \Delta t^{-m}$ under Assumptions 3 and 4. By substituting Eq.(3.35), we can obtain the following relation.

$$\begin{aligned} &\sum_{m=1}^{p-1} S_m \Delta t^{-m} = \\ &V \sum_{m=1}^{p-1} (-1)^m \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!} \Psi_j + \sum_{m=1}^{p-1} (-1)^{p+m-1} \frac{p!}{(p-m)!m!} I \Delta t^{-p}. \end{aligned} \tag{3.40}$$

To derive Ψ_0, we rewrite Eq.(3.12) by using Eq.(3.28) and $S_p = -I$ as follows.

$$\Psi_0 = I - V^{-1} \left(I \Delta t^{-p} - \sum_{m=1}^{p-1} \Delta t^{-m} S_m - S_0 \right). \tag{3.41}$$

By substituting Eq.(3.39) and Eq.(3.40) into this equation, we can obtain the following.

$$\begin{aligned}
&\Psi_0 = I - V^{-1}\Delta t^{-p} + \sum_{m=1}^{p-1}(-1)^m \sum_{j=m}^{p-1} \frac{j!}{(j-m)!m!}\Psi_j \\
&+V^{-1}\sum_{m=1}^{p-1}(-1)^{p+m-1}\frac{p!}{(p-m)!m!}I\Delta t^{-p} \\
&+V^{-1}\Bigg\{ I\Delta t^{-p} + (-1)^p\Psi_p^{-1}(I-\Psi_0)\Delta t^{-p} \\
&-\sum_{m=1}^{p-1}\left((-1)^m V\sum_{j=m}^{p-1}\frac{j!}{(j-m)!m!}\Psi_j + (-1)^{p+m-1}\frac{p!}{(p-m)!m!}I\Delta t^{-p}\right)\Bigg\} = \\
&I + V^{-1}(-1)^p\Psi_p^{-1}(I-\Psi_0)\Delta t^{-p}. \qquad (3.42)
\end{aligned}$$

From Eq.(3.7a), we see that $\Phi_p = (I-\Psi_0)^{-1}\Psi_p$, where $I-\Psi_0$ is always regular by Assumption 3 and Eq.(3.12). Since Ψ_p is regular by Eq.(3.36) and Assumption 3, Φ_p is also regular. Thus, we can write as $\Phi_p^{-1} = \Psi_p^{-1}(I-\Psi_0)$. By substituting it into Eq.(3.42) together with Eq.(3.32), we derive the following expression.

$$\begin{aligned}
\Psi_0 \;&=\; I - V^{-1}(-1)^p\Phi_p^{-1}\Delta t^{-p} = \\
&\quad I + (-1)^p\Delta t^{-p}\left(I - \sum_{k=1}^{q}(-1)^k\sum_{j=k}^{q}\frac{j!}{(j-k)!k!}\Omega_j\right)\Phi_p^{-1}.
\end{aligned}$$

We change the orders of the double summations with Ω_j as follows.

$$\begin{aligned}
&\sum_{k=1}^{q}(-1)^k\sum_{j=k}^{q}\frac{j!}{(j-k)!k!}\Omega_j = \sum_{j=1}^{q}\Omega_j\sum_{k=1}^{j}(-1)^k\frac{j!}{(j-k)!k!}. \\
&\therefore \Psi_0 = I + (-1)^p\Delta t^{-p}\left(I - \sum_{j=1}^{q}\Omega_j\sum_{k=1}^{j}(-1)^k\frac{j!}{(j-k)!k!}\right)\Phi_p^{-1}. \qquad (3.43)
\end{aligned}$$

By substituting Eq.(3.8b), we eliminate Ω_j from Eq.(3.43) and further reorganize as follows.

$$I - \Psi_0 = (-1)^p\Delta t^{-p}\left(\sum_{j=1}^{q}(I-\Psi_0)\Theta_j(I-\Psi_0)^{-1}\sum_{k=1}^{j}(-1)^k\frac{j!}{(j-k)!k!} - I\right)\Phi_p^{-1}.$$

Multiplying by $\Phi_p(I-\Psi_0)$, we can obtain the following.

$$(I-\Psi_0)\Phi_p(I-\Psi_0) = \\ (-1)^p \Delta t^{-p} \left((I-\Psi_0) \sum_{j=1}^{q} \Theta_j \sum_{k=1}^{j} (-1)^k \frac{j!}{(j-k)!k!} - (I-\Psi_0) \right).$$

By further multiplying $\Phi_p^{-1}(I-\Psi_0)^{-1}$, we can derive Eq.(3.18). $\square$

Chapter 4

Analysis of BWR instability mechanisms applying a CSARMA approach

4.1 Introduction

The development and/or application of time-series-analysis (TSA) and system identification techniques dedicated to instrumentation, monitoring and diagnostics of Light-Water-Reactor (LWR) processes or for inference of physical mechanisms from the latter, has been and continues to be a matter of intensive research (Williams, 1977; Fry et al., 1984; Sweeney, 1987; Pasczit, 1999; Zylbersztejn et al., 2013). For Boiling-Water-Reactors (BWRs), a specific and important target application of TSA that can be mentioned is the evaluation of the reactor stability properties using measured neutron flux signals as basis.

The BWR constitutes indeed a dynamical system and can as such be considered as stable when returning to an asymptotically equilibrium condition after that an initial departure from this condition has occurred due to some disturbances. But in this context, one must distinguish between core stability, which refers to fluctuations in coolant flow and power generation process coupled via nuclear feedback, and plant stability which instead refers to interactions of control and mechanical systems (*e.g.*, reactor pressure, feedwater flow and pump) with processes occurring in the core. To evaluate the core stability properties, system identification methods based on univariate AutoRegressive Moving-Average (ARMA) modeling of neutron flux signals are usually applied (Rotander, 1999). This is for instance

the case in the Paul Scherrer Institut (PSI) methodology developed over the years for applications to the Swiss reactors (Dokhane et al., 2006). Typically, the ARMA models are estimated based on the assumption that process noise (*e.g.*, pressure and temperatures) excites the core dynamics and that the driving noise is white. The advantage is that the ARMA models can then be used in a black-box manner when estimating the core stability properties, namely the decay ratio (DR) and the resonance frequency (RF), even if model parameters have no physical correspondence. Now regarding oscillatory modes related to plant stability, these can in most case be clearly separated from the core stability mode. However, interactions between the two could under certain circumstances take place and in this case, it might no longer be straightforward to identify the noise source that is driving the fluctuations and to establish thus a cause and effect relationship (Kwatny & Fink, 1975; Wilson, 2006; Kanemoto et al., 1982).

This was precisely the case for an event that occurred in recent years in a Swiss BWR plant (Ferroukhi, 2008). This event was found to have been caused by intense interactions between undamped neutron flux oscillations and specific process signal fluctuations but without revealing the causality relationship.

To address this, causality analysis methods based on multivariate models, *e.g.*, the DVAR and the DARMA models, become necessary (Ferroukhi, 2008; Oguma, 1981). For the mentioned above analysis methods, the parameters of such DVAR/ DARMA model do not need to have any physical correspondence. Also, these DVAR/DARMA models rely on discreet-time representations of the system dynamics. However, as mentioned in previous chapters DVAR/DARMA models do not uniquely represent the objective reactor processes (Gottschalk, 2001; Moneta et al., 2010; Demeshko et al., 2014). This limitation makes difficult to understand the dynamics of the objective system and its individual processes. In this chapter, we apply CSARMA modeling method, that is precisely aimed to overcome this limitation. This approach derives canonical CARMA and SARMA models of the systems dynamics that uniquely represent the system and where each equation in the models has bijective correspondence to each individual process in the system. These models allow us to investigate the dependency among the processes in the system and give us more insights on the physical mechanisms of the system (Moneta et al.,

2010; Stamer et al., 1996). Through this, the canonical models matrices should then indicate physical relationships between the signals and this could in turn allow to study underlying interactions and/or gain deeper insight into physical mechanisms of the system dynamics.

So far the verification of the CSARMA method has been limited to numerical cases and it therefore was considered valuable to attempt assessing its applicability for reactor noise analysis in general and for causality analysis in particular. To that aim, the KKL plant instability event was considered as appropriate situation target and this is thus the scope of the work presented in the following sections. Section 4.2 gives the description of analyzed KKL plant instability event. Section 4.3 shows the results of CSARMA applications to the KKL Cycle 24. The description and the results of STP method application to the KKL Cycle 24 are presented in Section 4.4. The discussion of the two methods applications is made in Section 4.5. Section 4.6 presents our conclusions.

4.2 KKL plant instability event

4.2.1 Description

During start-up of a 24th KKL cycle, high decay ratios (DR) with a corresponding large resonance frequency (RF) were suddenly indicated by the plant on-line stability monitoring system. During that time, the reactor was in the power ascension phase with the specific objective to raise the thermal power from 18% to 27% while maintaining a core flow around 34%[1]. The time evolution of some main process signals preceding, during and following the event is illustrated in Fig. 4.1. For simplicity, the time-frame has been decomposed into 6 periods with the main characteristics summarized in Table 2.1.1. The core was thus indicated as having become unstable during period 4 and 5 which lasted each around 10 minutes. However, during the entire time frame, the power/flow ratio (P/F) remained below 0.65 which is clearly very low compared to the P/F ratio range where core instability mechanisms might potentially be triggered.

[1] 100% power = 3600 MW, 100% flow = 11151 kg/s

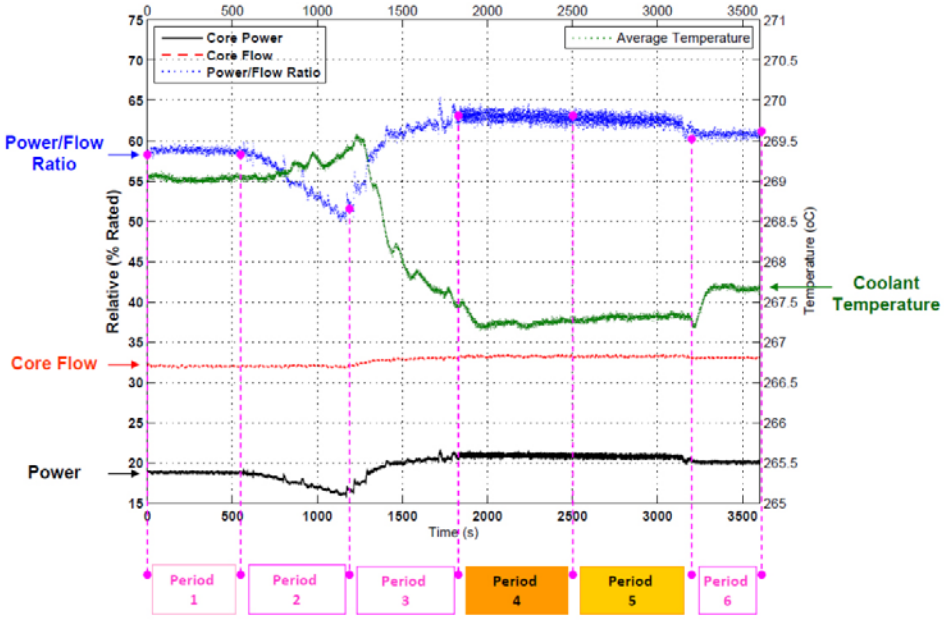

Figure 4.1: Time evolution of main process parameters during KKL plant instability event.

4.2.2 Time series analysis

Because of the unexpected occurrence of undamped neutron flux oscillations in this operating range, a time-series-analysis (TSA) of available measured signals was conducted shortly after the event (Ferroukhi, 2008). On the one hand, one objective was to verify if the instability alarms were triggered by erroneous evaluations of the core stability properties by the on-line stability monitoring system. The reason is that the latter employed a non-parametric method based on autocorrelation functions (ACF) which do not contain any phase information. Thereby, in case of several underlying oscillations at different frequencies, the ACF results might be associated with high uncertainties depending on the relative differences in power spectral density of the various oscillation components. Therefore, the PSI methodology based on ARMA model identification from measured neutron flux signals (Wilson, 2006) was applied and the results are illustrated in Fig. 4.2, noting that results shown there were estimated based on a weighting of the underlying dominant oscillation frequencies upon their associated power spectral density. Clearly,

Table 4.1: Characteristics of time periods during KKL plant instability event.

Period	Process Characteristics	Reactor Manoeuvres	On-Line Stability Monitoring
1	Stationary	—	Stable
2	Non-Stationary	Power Reduction (CR insertion and reduced Subcooling)	Stable
3	Non-Stationary	Power Increase (Feedwater Injection followed by CR withdrawal)	Stable
4	Stationary	—	Unstable
5	Stationary	—	Unstable
6	Stationary	None but following a short power reduction (CR Movements and reduced Subcooling)	Stable

this analysis confirmed the correctness of the on-line stability indications with a jump in frequency up to 0.9 Hz accompanied by a decay ratio close to 1.0.

To understand the reasons for the above behavior, the TSA was enlarged to a complete signal analysis of available measured signals including thus spectral, phase and coherence analyses. For simplicity, results will hereinafter only be reported for Periods 1 and 6 to represent time-frames for stable conditions prior and after the event respectively. And only results for Period 4 will be reported as representative of the unstable event time frames. Hence, for the stable Periods 1/6 and the unstable period 4, the power spectral density of neuron flux and selected process signals are shown in Fig. 4.3. Clearly during Period 4, a strong spectral peak previously not present appeared at around 0.9-1 Hz in all the signals with marquantly high spectral density especially in the steam flow signals. Once the core went back to indicated stable conditions (Period 6), such peak remained in the steam flow and neutron flux although with much lower power density and basically disappeared from the reactor pressure spectrum.

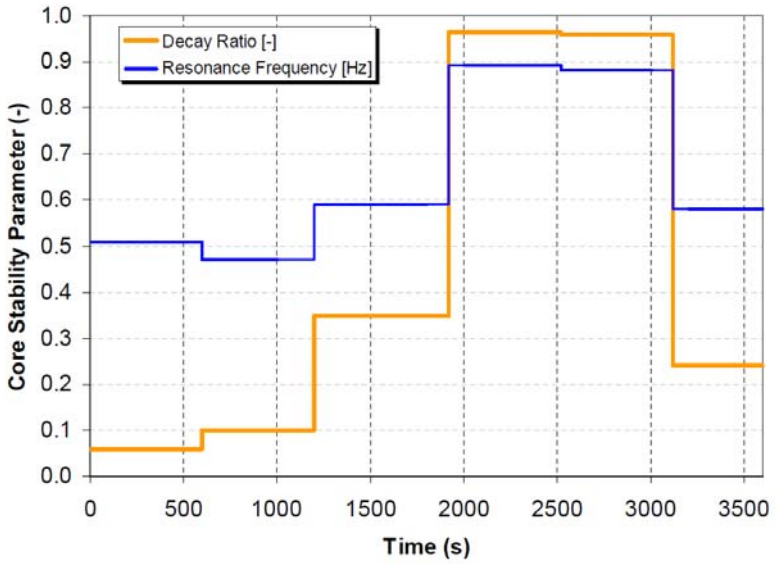

Figure 4.2: PSI evaluation of decay ratio and resonance frequency during instability event.

To further investigate the observed peaks, phase- and coherence analyses between all signals were for all periods conducted. For Period 4, illustrative results of the coherence analysis between neutron flux and steam flow are illustrated in Fig. 4.4. There, corresponding results from a stability test carried out at the plant and during which the core was brought close to unstable conditions, are also shown. For this test conducted at very low flow conditions, the measured decay ratio and resonance frequency evaluated from neutron noise signals were 0.97 and 0.47 Hz respectively.

From Fig. 4.4, the first observation is a full and COH between neutron flux and steam flow around 0.9-1 Hz. The same trend was observed between neutron fluxes and all other signals as well as between the latter. Thereby, it appeared clear that the undamped neutron flux oscillations causing the instability alarm were clearly the result of interactions with process signals related to pressure control (*e.g.*, via steam flow) and/or eventually feedwater control. As causality analysis methods were not implemented at the time of this TSA, the comparison with the stability test served the purpose of providing indications of the relationship between (core) neutron flux and (plant) steam flow. Because in this case, it was known that the core

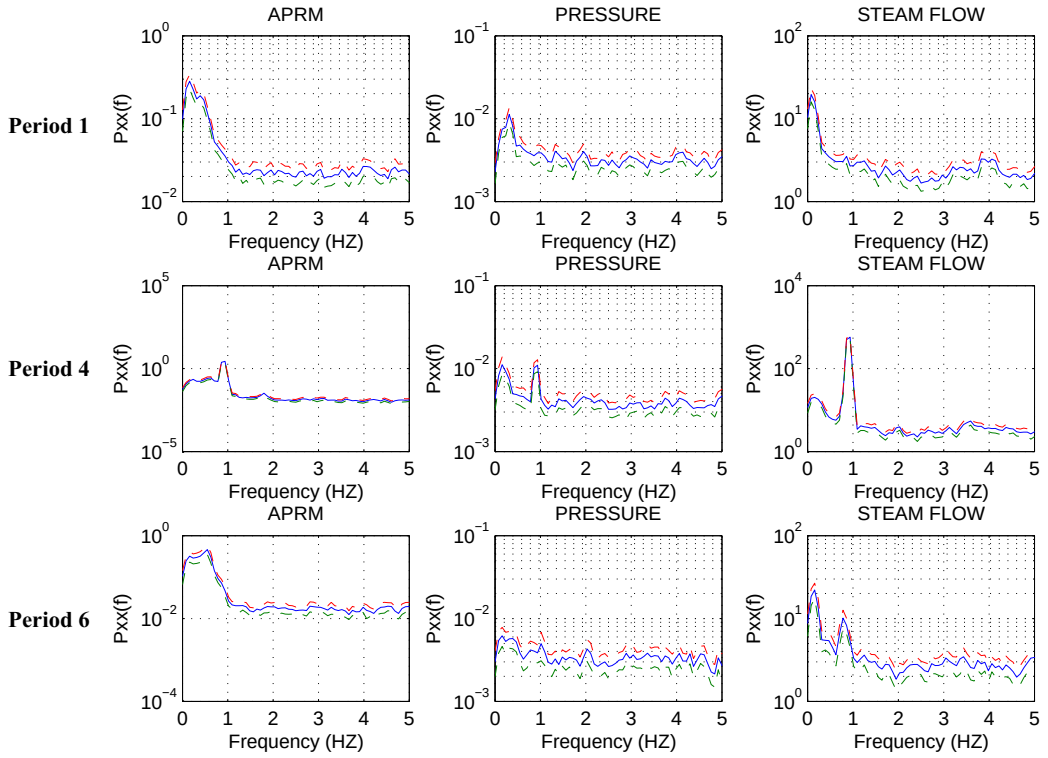

Figure 4.3: Spectral analysis of selected measured neutron and process signals.

was close to unstable without any interactions from plant control systems. This is indeed reflected by a very distinct large neutron flux spectral peak at around 0.5 Hz, *i.e.*, in the expected core natural frequency range while the stream flow only shows a slightly damped peak above this frequency, something that could to some extent be indicative of a shift due to the time constants between fission power oscillations and resulting steam outflow response. However, the main observation deduced from the practically zero COH was that in case of an unstable core with strong neutron flux oscillations, this would not produce corresponding strong oscillations of the steam flow. Thereby, it was concluded that most likely, the start-up stability event was driven by noise disturbances and fluctuations stemming from interactions with plant control systems. Recently, work towards confirming this by causality analysis methods was initiated and this is thus presented in the flowing sections.

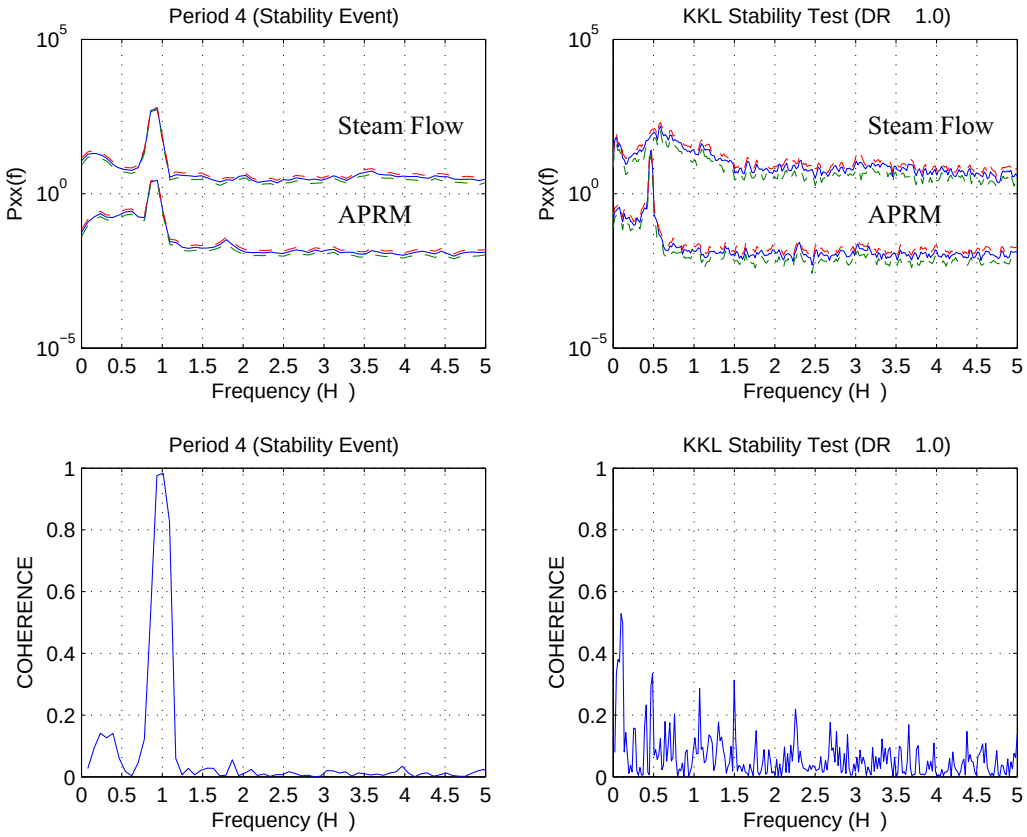

Figure 4.4: Comparison of spectral- and coherence analysis for neutron flux and steam flow between KKL plant instability event and stability test with close to unstable core.

4.3 Application of CSARMA for causality analysis of KKL plant instability event

For understanding the physical mechanisms of the reactor dynamics, we need to have a mathematical model and its corresponding parameters that precisely reflect the reactor processes. Since the reactor noise generation processes can be often approximated by a multivariate linear Markov system, we apply the developed CSARMA modeling approach. It derives SARMA model that represents the

physical dynamics of the system in canonical manner, and the parameter matrices Ψ_j and Ω_j of the SARMA model give us the information on the fundamental processes in the system. Finally, by using the relations between the SARMA and the CARMA models presented in Eq.(3.15)-(3.17), we estimate parameters of the CARMA model of the system (Demeshko et al., 2014).

For the analysis of the KKL plant instability event, we point out that particular intense interactions between reactor pressure (PRESS), neutron fluxes (LPRM) and steam flow (SFLOW) were shown by the time series analysis explained in Section 2. For these reasons, the CSARMA method is applied here for a 3-dimensional multi-variate system comprising $Y = [PRESS, LPRM, SFLOW]^t$.

For the CSARMA application, it is necessary to know the model orders of the system. They were provided by HPTSAC method developed at Paul Scherrer Institute, Switzerland (Dokhane, 2004), based on the combination of Akaike information criterion, Plateau method and Minimum description length principle. According to the HPTSAC estimations, the dynamics of the objective processes might be approximately represented by DARMA(4,3) model. Accordingly, for each period, we applied CSARMA method deriving CARMA (4,3) model from the estimated DARMA(4,3) model, where the orders of the CARMA model correspond to the orders of differential equations representing the system.

The CARMA autoregressive matrices S_m for m=0,1,2,3 and moving-average matrices R_n for n=1,2,3 represent the dependency mechanisms among variables, where the first column shows the effects of reactor pressure on the other variables, the second column - the effects of LPRM and the third one - the effects of steam flow. The (i,j)-element of each matrix represents the effect from the j-th variable to the i-th in the variable vector Y, and its absolute value represents the intensity of the effect, *e.g.*, (1,2)-element represents the effect from LPRM to reactor pressure.

For periods 1,4 and 6 of KKL plant instability event, we first evaluated the ratios of the average value of the MA matrices over that of the AR matrices as follows:

$$\varphi = \frac{\frac{1}{q}\sum_{k=1}^{q}\frac{1}{d^2}\sum_{i,j}|r_{k,ij}|}{\frac{1}{p}\sum_{k=1}^{p}\frac{1}{d^2}\sum_{i,j}|s_{k,ij}|},$$

where $|r_{k,ij}|$ and $|s_{k,ij}|$ are absolute values of i,j-elements for k-order MA and AR matrices, respectively. The ratios are 0.14%, 0.18% and 0.19% for periods 1,4 and

6, respectively. In addition, there are no exceptionally large elements in any MA matrices that do not follow these ratios. Thus, we exclude the MA part of the CARMA models in our analysis and concentrate on the AR part.

According to the above notation, the estimated diagonal elements of the CARMA matrices represent the effect of a signal on itself, *i.e.*, "self-effects". They are shown for the various signals, periods and matrices in Fig. 4.5, from which a few observations can be made. First, for each signal, the self-effects show a cyclic nature (change of sign) as function of matrix order, *i.e.*, time shift. Secondly, for any matrix order, it is interesting to note a relatively much stronger self-effect for the steam flow compared to the two other signals. This might indicate that the steam flow is influenced by some external process not included in the three-signals system investigated here. That might be a disturbance in some control system or the influence of the feed water system, which can effect steam flow on such low power levels (Kwatny & Fink, 1975). However, these are just hypotheses and they need to be clarified in the future by additional investigation. Third, it is noted that the two stable Periods 1/6 show a very similar behavior. However, also Period 4 show a similar pattern, the only minor difference being a reduction of the LPRM and steam-flow self-effects combined with an increase of the pressure self-effect. However, overall, the patterns are rather similar between all periods indicating that the self-effects, *i.e.*, the CARMA diagonal terms, do not provide enough information to allow discriminating the core/plant dynamics between Periods 1/6 and 4.

Next, our focus is given to the effects between variables, *i.e.*, the off-diagonal elements of the CARMA matrices. The results of the CARMA model excluding diagonal matrices are shown in Fig. 4.6. The following three major observations are obtained based on the qualitative features indicated by the CARMA matrix off-diagonal elements.

- First, the causal structure is quite similar in the stable periods 1 and 6. This indicates that the core/plant dynamics represented by the CARMA model returned to the similar dynamics after the instabilities occurred in period 4.

- Secondly, the less stable period 4 shows on the other hand distinctly different patterns compared to the two other periods. This indicates that the CARMA

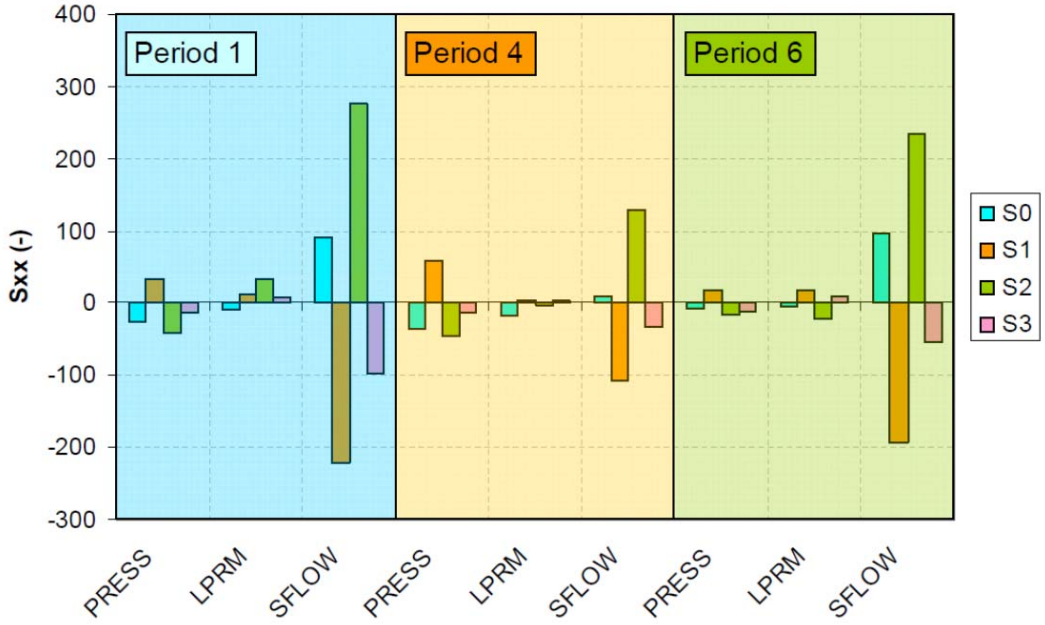

Figure 4.5: CSARMA results for self-effects - diagonal matrix elements.

model has captured a complete change in the dynamical interactions between the three investigated variables during this period. And this is fully in-line with the fact that among all three periods, a core/plant related instabilities occurred indeed only during period 4.

- Third, we note that the consistency between period 1/6 and the distinctly difference of period 4 are observed in every matrix order, *i.e.*, each time differential in the dynamics between the variables.

By closely looking at this figure, we notice that the dominant effects, {SFLOW $\Longrightarrow$ PRESS}, {LPRM $\Longrightarrow$ PRESS} and {LPRM $\Longrightarrow$ SFLOW} are observed in some orders of the AR matrices at periods 1/6. For period 4, a radically different behavior is seen. Indeed, the dominant effect is {PRESS $\Longrightarrow$ SFLOW } only, however we also observe some minor effects like {PRESS $\Longrightarrow$ LPRM}, {SFLOW $\Longrightarrow$ LPRM} and the feedback {SFLOW $\Longrightarrow$ PRESS} additionally. In other words, PRESS and SFLOW have bidirectional causality while the influence from PRESS

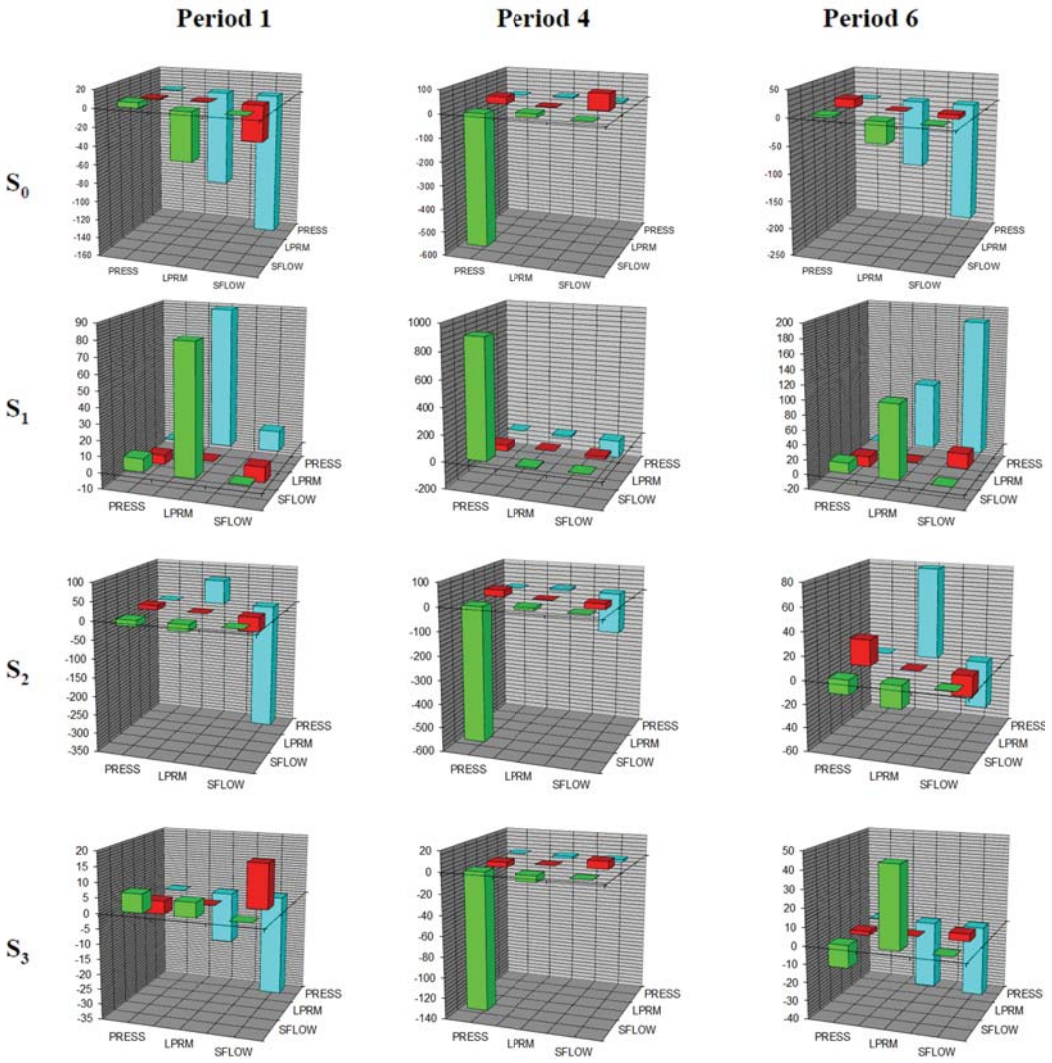

Figure 4.6: CSARMA results for casual structure - off-diagonal elements, where the blue bars show the elements of the first row, the red bars - the elements of the second row, and the green bars - the elements of the third row of CARMA models matrices.

is much bigger than other way around, which is radically differs from periods 1 and 6. Also the CSARMA results for period 4 indicate, it was {PRESS $\Longrightarrow$ SFLOW } disturbances that induced/guided LPRM fluctuations and not the other way around. Finally, it is worth noting that the changes of the signs of elements show

some oscillatory behavior over matrices orders for all the effects.

4.4 Benchmarking against STP method

4.4.1 Signal transmission path analysis

The signal transmission path analysis has been widely used in the causal relationship analysis by the aid of multivariate time series modeling (Oguma, 1981, 1982a, 1982b; Oguma & Turkan, 1985; Oguma, 1996). In addition, using this analysis allows detecting the feedback effect among variables of interest and allowing information about how the noise power is propagating from one variable to another.

In the current study, the STP analysis is performed for pairs of variables of interest, using the main two functions, *i.e.*, the coherence (COH) and the noise power contribution (NPC). The coherence function is very important to be evaluated since it represents the correlation between two variables in frequency domain and allows characterizing the dynamic relationship between the measured variables. The NPC function is used to evaluate noise source power propagation from one variable to another, allowing determination of the causal relationship between the variables of interest. The followings are the definitions of these two main functions (Oguma, 1981).

Consider a linear dynamic system with feedback effect between variables X_1 and X_2 that can be described as:

$$X_1(z) = G_{12}(z)X_2(z) + N_1(z), \tag{4.1}$$

$$X_2(z) = G_{21}(z)X_1(z) + N_2(z), \tag{4.2}$$

where $G_{21}(z)$ and $G_{12}(z)$ are the open loop transfer functions from X_1 to X_2 and from X_2 to X_1, respectively. $N_1(z)$ and $N_2(z)$ are noise sources to X_1 and X_2, respectively, and they are assumed to be stationary random with zero mean and mutually independent.

The auto power spectral density (APSD) functions $P_{11}(\omega)$ and $P_{22}(\omega)$ for the corresponding variables and the cross power spectral density (CPSD) function $P_{12}(\omega)$

are given by:

$$P_{11}(\omega) = \left({}^{1}/_{\Delta t^2}\right)\left(|G_{12}(\omega)|^2 Q_{22}(\omega) + Q_{11}(\omega)\right), \tag{4.3}$$

$$P_{22}(\omega) = \left({}^{1}/_{\Delta t^2}\right)\left(|G_{21}(\omega)|^2 Q_{11}(\omega) + Q_{22}(\omega)\right), \tag{4.4}$$

$$P_{12}(\omega) = \left({}^{1}/_{\Delta t^2}\right)\left(G^{*}_{21}(\omega)\, Q_{11}(\omega) + G_{12}(\omega)\, Q_{22}(\omega)\right), \tag{4.5}$$

where $\Delta = |1 - G_{12}(\omega)G_{21}(\omega)|$, $Q_{11}(\omega)$ and $Q_{22}(\omega)$ are the APSDs of the respective noise sources, $|\cdot|$ denotes the absolute value and the superscript * the complex conjugate. Using Eq. (4.3)-(4.5), we define the COH function $\gamma^2_{12}(\omega)$ between X_1 and X_2 as

$$\gamma^2_{12}(\omega) = \frac{|P_{12}(\omega)|^2}{P_{11}(\omega)\, P_{22}(\omega)}. \tag{4.6}$$

The NPC ratio from X_2 to X_1 through the noise source N_2 is

$$\Gamma_{12}(\omega) = \frac{|G_{12}(\omega)|^2 Q_{22}(\omega)}{|G_{12}(\omega)|^2 Q_{22}(\omega) + Q_{11}(\omega)}. \tag{4.7}$$

Similarly, the NPC ratio from X_1 to X_2 through the noise source N_1 is

$$\Gamma_{21}(\omega) = \frac{|G_{21}(\omega)|^2 Q_{11}(\omega)}{|G_{21}(\omega)|^2 Q_{11}(\omega) + Q_{22}(\omega)}. \tag{4.8}$$

Note that the mutual comparison of COH and NPC functions allows the identification of causal relationship and the detection of the feedback between the two variables. Hence the following possibilities may exist (Oguma, 1996):

- The two functions are zero at all frequencies if no dynamic relationship between the two variables exists.
- If there is an explicit causal relationship only from X_1 to X_2, *i.e.*, no feedback, then the NPC function from X_1 to X_2 coincide with the COH function, while the NPC function from X_2 to X_1 is zero at all frequencies, and vice versa.
- The two functions differs from each other if there is a feedback between the two variables.

It should be emphasized that the current STP analysis is carried out based on the DARMA modeling technique using the NAG subroutine g13dc. Under the condition

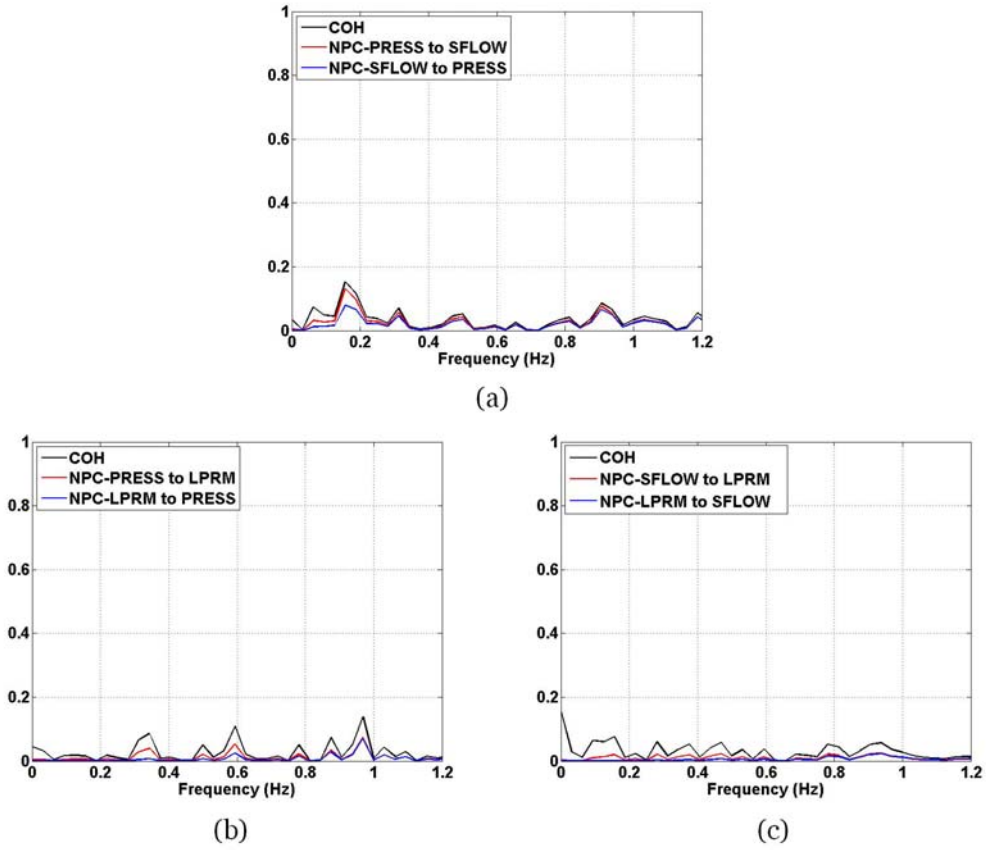

Figure 4.7: Application of STP to KKL plant instability event period 1.

of independent noise sources between the pair variables, the transfer functions and the power spectra of the noise sources are derived in terms of the DARMA model parameters. Because STP analysis is based on non-canonical DARMA model, its results can have some discrepancy with the causal structure of the actual nuclear plant.

4.4.2 Application to KKL plant instability event

Similar to CSARMA, the STP method is applied to each of the three periods, *i.e.*, 1, 4 and 6. The results for periods 1 and 6 are similar, therefore, for shortness we show only the one for period 1 in Fig.4.7. The results for period 4 are shown in

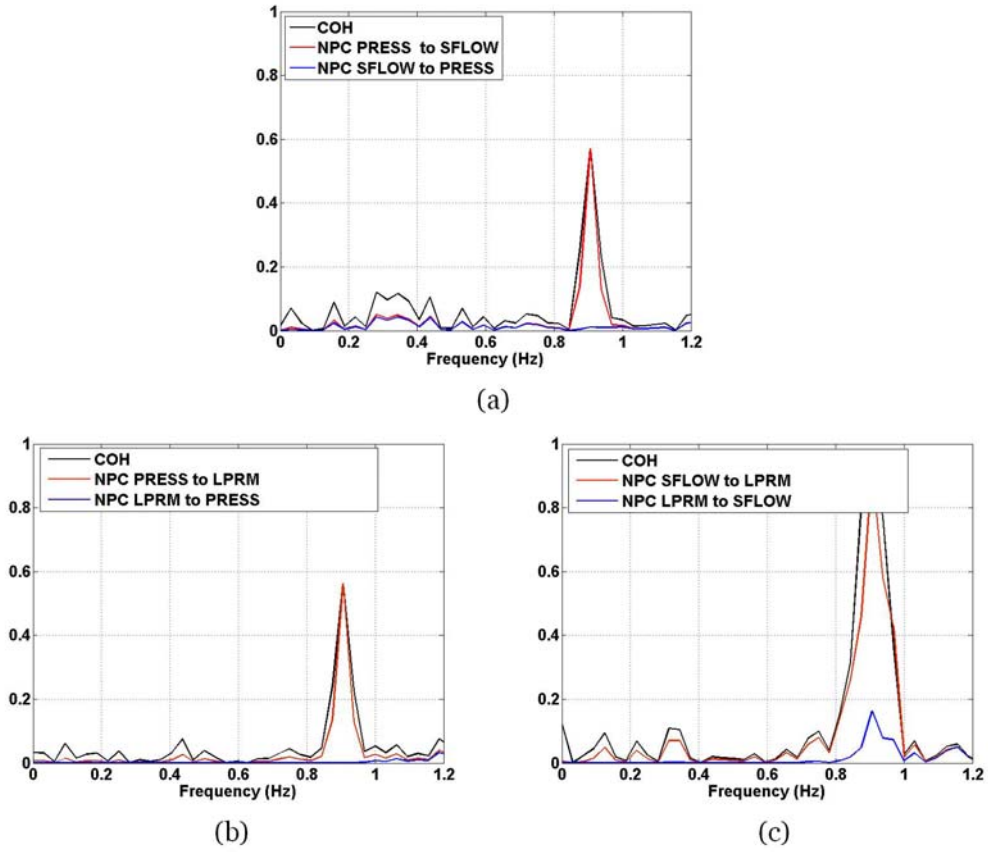

Figure 4.8: Application of STP to KKL plant instability event period 4.

Fig.4.8. We see that for period 1 and 6, the COH and NPC functions for all signal pairs are almost less than 0.1, which is negligible. This means the STP method does not show any strong dynamic relationship among all the three variables, *i.e.*, steam flow, reactor pressure and LPRM. However, for period 4, as expected (see section 2.2), a clear peak at around 0.9 Hz is observed for the COH function of all the signal pairs with a maximum value for the pair (LPRM, steam flow), indicating a COH between the neutron flux and the steam flow at 0.9 Hz. As can be seen from Fig.4.8, the NPC from reactor pressure to steam flow and neutron flux is dominant, while the NPC in the opposite direction, *i.e.*, from neutron flux to reactor pressure and from steam flow to reactor pressure, is negligible. In addition, the NPC from steam

flow to neutron flux is dominant while it is negligible in the opposite direction.

4.5 Discussion

Here, we note that the STP method failed to indicate any effects among variables for periods 1 and 6. However, the STP results for period 4 are in total agreement with those obtained using CSARMA approach.

Also, the CSARMA method results for periods 1 and 6 indicated several dominant effects. The dominance of {SFLOW $\Longrightarrow$ PRESS} reflects correctly the "reactor master-turbine slave" concept with the pressure controller adjusting the opening of the turbine inlet/bypass valves to regulate the reactor pressure. The {LPRM $\Longrightarrow$ PRESS} dominance also follows from the same concept, *i.e.*, during normal operation without any ex-vessel or system induced disturbances, the reactor pressure will be held constant and will increase/decrease as a response to a sudden core power (LPRM) increase/decrease. Hence, in this mode, the neutron flux will guide the reactor pressure evolution through the pressure controller, also guide the steam flow behavior as reflected by the observed {LPRM $\Longrightarrow$ SFLOW } dominance.

For period 4, both CSARMA and STP analysis methods indicated the opposite behavior of the influence from reactor pressure to steam flow which is inconsistent with the conventional "reactor master-turbine slave" concept. Thus, we can assume that reactor pressure is driven by some mechanisms which are not included into the "reactor master-turbine slave" concept. This would be in-line with the expected physics of the origin of the event, *e.g.*, a fluctuation at the turbine/bypass valve inlet zone that propagated back to the vessel, inducing disturbances in the pressure control system.

4.6 Conclusions

This chapter has presented a first application of a novel Continuous and Structural Autoregressive Moving Average (CSARMA) modeling approach to BWR noise analysis. This method derives more robust and reliable models as basis for signal analysis in general and for reactor diagnostics or causality analysis in particular. As a first step towards assessing the potential of CSARMA the latter type of applications, a stability event that occurred in a Swiss BWR plant during power ascension phase was analyzed. To that aim, CARMA models for a multi-variate system comprising steam flow, reactor pressure and LPRM were estimated based on measurements taken during the closed to instability event. Similarly, corresponding CARMA models were also estimated for more stable time periods that preceded and followed the event.

On that basis and focusing only on qualitative results at this stage, the estimated time-domain based CARMA matrix structures were studied as function of model order as well as between the various selected time periods. As main observations, it was found that the CARMA matrix structures would clearly indicate a different dynamical state during the less stable period 4 compared to the two more stable periods 1 and 6. Moreover, during the period 4, the CSARMA results would show a distinctly dominant relationship from steam flow to neutron flux and not vice versa, indicating thus that the core instability was caused by a plant system disturbance rather than the opposite. As well, a very dominant relationship from reactor pressure to steam flow was indicated, this being in total opposition to the normal situation during reactor master/turbine slave operation when reactor pressure is adjusted through the steam flow. Hence, these CSARMA results could be interpreted as pointing out a disturbance in the pressure control system as primary cause for the instabilities during period 4.

To benchmark all these findings, the frequency-domain based Signal Transmission Path (STP) method, currently being implemented at PSI as part of the overall noise analysis methodology for the Swiss reactors, was also applied. And with the STP method, for period 4 exactly the same relationships as mentioned above were obtained. On the one hand, this consistency between both methods could

be considered as having confirmed that the period 4 instabilities were caused by a steam flow disturbance and not induced by the core. On the other hand, this consistency could also be considered as a verification of the CSARMA ability to correctly indicate the causality relationships even during less stable event.

It is also worth to note that for periods 1 and 6 the STP analysis failed to catch the dynamical relations between variables. The possible reason for it might be that the STP analysis is based on non-canonical DARMA model, therefore, its results can have some discrepancy with the causal structure of the actual nuclear plant. Also, the STP analysis is usually unable to estimate the interactions among variables for the cases when the system has weak noise function, *e.g.*, very stable cases. However, the CSARMA modeling approach clearly indicated intense effects from both steam flow and neutron flux to reactor pressure for periods 1 and 6. These results fully correspond to the background knowledge on the BWR physics.

Chapter 5

Conclusion

This dissertation investigated new modeling approaches to derive canonical VAR and ARMA models. In particular, we considered structural and continuous time VAR and ARMA models, as the basis of our framework.

First, we worked on approach called CSVAR method. It assumes that the objective system is continuous time, multivariate, linear, Markov, stable and controllable system. This approach is based on the mathematical relations between the CVAR, SVAR and DVAR models. The advantages of CSVAR modeling over the existing modeling approaches have been verified numerically.

Next, we considered a modeling approach advanced to ARMA processes, what we called CSARMA method. Upon on the general assumptions on the system, it allows us to derive the canonical CARMA and SARMA models from the DARMA model estimated by using time series data observed from the system. Both the CARMA and the SARMA models are general representations of a continuous time, multivariate, linear Markov system in comparison with the CVAR and SVAR, respectively. The advantages of CSARMA modeling over the past modeling approaches have been verified through numerical experiments.

Last, we applied the developed CSARMA modeling approach to the real world system of BWR for the investigation of the occurred instability event. This application showed the superior performance of our method over the STP method, traditionally used for nuclear reactor analysis. It gave us dynamical relations between variables, which STP method failed to indicate and allowed us to conclude that a primary cause for the instabilities observed during reactor operation was a disturbance in the pressure control system.

Apart from the remaining problems raised in each chapter, we point out a general

issue as the further improvement direction of this study. While CSARMA aims at overcoming a limitation of time domain methods in terms of ensuring a unique representation of the system dynamics, its application via time-domain structural matrices to infer causality relationships might not be sufficient. A transformation of the latter to the frequency domain could therefore become necessary and is therefore planned as part of further method developments and verifications of this method.

References

Blomstrand, J., Rotander, C., & Lundberg, S. (1993, 25-27 May). Core stability monitoring. on-line performance qualification of COSMOS in KKL. In Proceedings Jahrestagung Kerntechnik (pp. 361-364). Cologne, Germany.

Brockwell, A. & Brockwell, P. (1999). A class of non-embeddable arma processes. Journal of Time Series Analysis, 20 (5), 483-486. doi: 10.1111/1467-9892.00151.

Brockwell, P. & Davis, R. (1991). Time Series: Theory and Methods (Second edition ed.). New YorkSpringer.

Brockwell, P. & Peter, J. (1995). A note on the embedding of discrete-time arma processes. Journal of Time Series Analysis, 16 (5), 451-460.doi: 10.1111/j.1467-9892.1995.tb00246.x

Chambers, M. & Thornton, M. (2012). Discrete time representation of continuous time arma processes. Econometric Theory, 28 (01), 219-238.

Chen, Y., Zhao, H., & Yu, L. (2010). Demand forecasting in automotive aftermarket based on ARMA model. In International Conference on Management and Service Science 1-4.

Demeshko, M., Washio, T., & Kawahara, Y. (2013). A Novel Structural AR Modeling Approach for a Continuous Time Linear Markov System A novel structural AR modeling approach for a continuous time linear markov system. In The First IEEE ICDM Workshop on Causal Discovery 2013 (CD2013).

Demeshko, M., Washio, T., Kawahara, Y., & Shimizu, S. (2014). A novel structural ARMA modeling across continuous and discrete time domains. In press.

Dokhane, A. (2004). BWR Stability and Biffurcation Analysis using a Novel Reduced Order Model and the System Code RAMONA. Unpublished doctoral dissertation, EPFL, Switzeland.

Dokhane, A., Ferroukhi, H., Zimmermann, M., & Aguirre, C. (2006). Spatial and Model-order Based Reactor Signal Analysis Methodology for BWR Core Stability Evaluation. Annals of Nuclear Energy, 33, 1329-1338.

Ferroukhi, H. (2008). Reactor Signal Analysis of KKL BOC24 Data Measured on the 22.08.2007 (Tech. Rep.). PSI.

Fisher, F. (1970). A correspondence principle for simultaneous equation models. Econometrica, 38(1), 73-92.

Fry, D., Leuba, J.-M., & Seeney, F. (1984). Use of Neutron Noise for Diagnostics of In-Vessel Anomalies in Light Water Reactors (Tech. Rep.). ORNL Report TM-8774.

Gannabathula, P. (1988). Arma model estimation for eeg using canonical correlation analysis. In Engineering in Medicine and Biology Society, Proceedings of the Annual International Conference of the IEEE (Vol. 3, pp. 1204-1205).

Gill, P. E. & Murray, W. (1972). Quasi-newton methods for unconstrained optimization. Journal of the Institute of Mathematics and its Applications, 9, 91108.

Gottschalk, J. (2001). An Introduction into the SVAR Methodology: Identification, Interpretation and Limitations of SVAR models (No. 1072). Institute of World Economics, Kiel.

Hinrichsen, D. & Pritchard, A. (2005). Mathematical Systems Theory I, Modelling, State Space Analysis, Stability and Robustness (Vol. 48). Springer.

Hyvarinen, A., Shimizu, S., & Hoyer, P. (2008). Casual modeling combining instantaneous and lagged effects: an identifiable model based on non-gaussianity. In Proceedings of the 25th International Conference on Machine Learning (pp.424-431). Helsinki, Finland.

Iwasaki, Y. & Simon, H. (1994). Causality and model abstraction. Articial Intelligence, 67 , 143-194.

Kanemoto, S., Andoh, Y., Yamamoto, F., Kitamoto, K., & Nunome, K. (1982). Identification of Pressure Control System Dynamics in BWR Plant by Multivariate Autoregressive Modeling Technique. Journal of Nuclear Science and Technology, 19 (1), 58-68.

Kawahara, Y., Shimizu, S., & Washio, T. (2011). Analyzing relationships between ARMA processes based on non-gaussianity of external in uences. Neurocomputing, 74 (12-13), 2212-2221.

Kilian, L. (2011). Structural Vector Autoregressions. Edward Elger.

Kizilkaya, A. & Kayran, A. (2006). ARMA model parameter estimation based on the equivalent MA approach. Digital Signal Processing, 16 , 670-681.

Kuroda, Y., Uchida, H., & Ohmasa, Y. (1985). Multivariate ARMA modeling with spline functions for reactor noise. Progress in Nuclear Energy, 15 , 849-852.

Kwatny, H. & Fink, L. (1975). Acoustics, stability, and compensation in boiling

water reactor pressure control systems. IEEE Trans. Autm. Control , 1 , 727-739.

Lacerda, G., Spirtes, P., Ramsey, J., & Hoyer, P. (2008). Discovering cyclic causal models by independent components analysis. In Proceedings of the Twenty-Fourth Conference on Uncertainty in Artificial Intelligence.

Larsson, E., Mossberg, M., & Soderstrom, T. (2006). An overview of important practical aspects of continuous-time arma system identification. Circuits, Systems and Signal Processing, 25, 17-46.

Lax, P. & Richtmyer, R. (1956). Survey of the stability of linear finite difference equations. Communications on Pure and Applied Mathematics, 9, 267-293.

LeVeque, R. (2007). Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-state and Time-dependent Problems. SIAM e-books.

Levy, H. & Lessman, F. (1992). Finite Difference Equations. Dover.

Mahata, K. & Fu, M. (2007). On the indirect approaches for carma model identification. Automatica, 43 , 1457-1463.

Mainassara, Y. B. & Francq, C. (2009). Estimating structural VARMA models with uncorrelated but non-independent error terms. Journal of Multivariate Analysis, 102 (3), 496-505.

Moneta, A., Chlass, N., Entner, D., & Hoyer, P. (2011). Causal search in structural vector autoregressive models. In JMLR Workshop and Conference Proceedings, Causality in Time Series (Vol. 12, pp. 95-118).

Moneta, A., Entner, D., Hoyer, P., & Coad, A. (2010). Causal Inference by Independent Component Analysis with Applications to Micro- and Macroeconomic Data. Jena Economic Research Papers(No.2010-031), 3-4.

Mooij, J., Janzing, D., & Scholkopf, B. (2013). From ordinary differential equations to structural causal models: the deterministic case. In Proceedings of the 29th conference on Uncertainty in Artificial Intelligence. NAG. (2013). Nag toolbox for matlab [Computer software manual].

Oguma, R. (1981). Method to evaluate signal transmission paths in dynamic systems and its application to reactor noise analysis. J. Nucl. Sc. Tech, 18, 835-844.

Oguma, R. (1982a). Cohenerence analysis of systems with feedback and its application to bwr noise investigation. Prog. Nucl. Energy, 9, 137-148.

Oguma, R. (1982b). Extended partial and multiple coherence analyses and their

application to reactor noise inves-tigation. J. Nucl. Sc. Tech., 19, 543-554.

Oguma, R. (1996). Investigation of power oscillation mechanisms based on noise analysis at forsmark-1 bwr. Ann. Nucl. Energy, 23(6), 469-485.

Oguma, R. & Turkan, E. (1985). Application of an improved multivariable noise analysis method to investigation of PWR noise; signal transmission path analysis. Progress in Nuclear Energy, 15, 863-873.

Pasczit, I. (1999). Development of core diagnostic methods and their application at swedish bwrs and pwrs. Jour. Nucl. Sc.Tech., 36, 473-485.

Pearl, J. (2000). Causality: Models, Reasoning, and Inference. Cambridge University Press.

Pepyolyshev, Y. N. (1988). Spectral characteristics of power noise parameters and fluctuations of neutron reflectors of nuclear reactor IBR-2. Preprint, JINR, Dubna. ((in Russian))

Perrott, M. & Cohen, R. (1996). An efficient approach to ARMA modeling of biological systems with multiple inputs and delays. IEEE Transactions on Biomedical Engineering, 43, 1-14.

Pfaff, B. & Kronberg, T. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. Journal of statistical software, 27(4), 1-32.

Priestley, M. (1981). Spectral Analysis and Time Series (Vol. 1). Academic Press, London.

Rotander, C. (1999). Methods to identify stability properties for a thermohydraulic channel and from reactor noise measurements (Doctoral dissertation, Royal Institute of Technology). Lcesniate Thesis TITA-S3-REG-9901.

Shea, B. L. (1987). Estimation of multivariate time series. Journal of Time Series Analysis, 8 , 95-110.

Shimizu, S., Inazumi, T., Sogawa, Y., Hyvarinen, A., Kawahara, Y., Washio, T., & Bollen, K. (2011). Directlingam: A direct method for learning a linear nongaussian structural equation model. Journal of Machine Learning Research, 12 , 1225-1248.

Shittu, O. & Yaya, O. (2009). Measuring forecast performance of ARMA and ARFIMA models: An application to us dollar/uk pound foreign exchange rate. European Journal of Scientific Research, 32 (2), 167-176.

Soderstrom, T. (1991). Computing stochastic continuous-time models from arma

models. International Journal of Control , 53 (6), 1311-1326. doi: 10.1080/00207179108953677.

Stamer, O., Tweedie, R., & Brockwell, P. (1996). Existence and stability of continuous time threshold ARMA process. Statistica Sinica, 6 , 715-732.

Strikwerda, J. (1989). Finite Difference Schemes and Partial Differential Equations. LondonChapman and Hall.

Sweeney, F. (1987). Utility Guidelines for Reactor noise Analysis (Tech. Rep. No. NP-4970). EPRI.

Tomasson, H. (2011). Some Computational Aspects of Gaussian CARMA Modelling (Tech. Rep.). Institute for Advanced Studies, Vienna.

Tran, D. (1992). Generalization of the signal transmission path method for ARMA model in reactor noise analysis. Annals of Nuclear Energy, 19(6), 341-345.

Tran, D. (2003). Application of autoregressive moving average model in reactor noise analysis. Annals of Nuclear Energy, 20(12), 815-822.

Voortman, M., Dash, D., & Druzdzel, M. (2010). Learning why things change: The difference-based causality learner. In Proceedings of the 26th conference on Uncertainty in Artificial Intelligence.

Williams, M. (1977). Reactor noise - SMORN II. Progress in Nucl. Energy, 1, 73-804.

Wilson, D. H. (2006). Identifying potential interaction between bwr plant dynamics and power system electromechanical oscillations.

Zhang, L. & Xu, L. (2010). Forecasting of fluctuations and turning points of power demand in China based on the maximum entropy method and ARMA model. In the 5th International Conference on Critical Infrastructure (CRIS) (pp. 1-6).

Zhao, J., Xu, L., & Liu, L. (2007). Equipment fault forecasting based on arma model. In Proceedings of International Conference on Mechatronics and Automation (ICMA) 2007 (pp. 3514-3518).

Zylbersztejn, F., Tran, H., Pazsit, I., Demaziere, C., & Nylen, H. (2013). On the dependence of the noise amplitude on the correlation length of inlet temperature fluctuations in PWRs. Ann. Of. Nucl. Energy, 57, 134-141.

List of Publications

Journal Papers

- Demeshko, M. & Washio, T. & Kawahara, Y. & Pepyolyshev, Y. A Novel Continuous and Structural VAR Modeling Approach and its application to reactor noise analysis. ACM Transactions on Embedded Computing Systems, 2015 (To appear).

- Demeshko, M. & Dokhane, A. & Washio, T. & Ferroukhi, H. & Kawahara, Y. & Aguirre, C. Application of Continuous and Structural ARMA Modeling for Noise Analysis of a BWR Coupled Core and Plant Instability Event. Annals of Nuclear Energy, Vol.75, pp.645-657, September 2014.

Conference Proceedings

- Demeshko, M. & Washio, T. & Kawahara, Y. & Pepyolyshev, Y. Application of a Continuous Time Structural ARMA Modeling to Analysis of a Nuclear Reactor. In Proc. of the 28th Annual Conference of the Japanese Society for Artificial Intelligence, Matsuyama, Japan, 2G1-4, pp.1-4, May 2014 (in Japanese Domestic Conference).

- Demeshko, M. & Washio, T. & Kawahara, Y. A Novel Structural AR Modeling Approach for a Continuous Time Linear Markov System. In Proc. of 2013 IEEE 13th International Conference on Data Mining Workshops (ICDMW), TX, USA, pp.104-113, December 2013.

- Demeshko, M. & Washio, T. & Pepyolyshev, Y. Structural Analysis of IBR-2 Based on Continuous Time Canonicality. Transactions of the American Nuclear Society, Washington D.C., USA, Vol.109, pp.1445-1448, November 2013.

- Demeshko, M. & Washio, T. & Kawahara, Y. (2012). A Novel Structural ARMA Modeling Approach to Reactor Noise Analysis. In Proc. of Reactor Noise Knowledge Transfer Meeting (RNKTM2012), Prague, Czech Republic, ISSN 1805-6156, paper ID 5, October 2012.

- Demeshko, M. & Washio, T. & Kawahara, Y. & Shimizu, S. Analyzing Relationships Between CTARMA and ARMA Models. In Proc. of the 25th Annual Conference of the Japanese Society for Artificial Intelligence, Morioka, Japan, 2G2-2, pp.1-4, June 2011 (in Japanese Domestic Conference).