Master Thesis (45 EC)

Frequency domain identification in dynamic networks for In-Circuit Testing

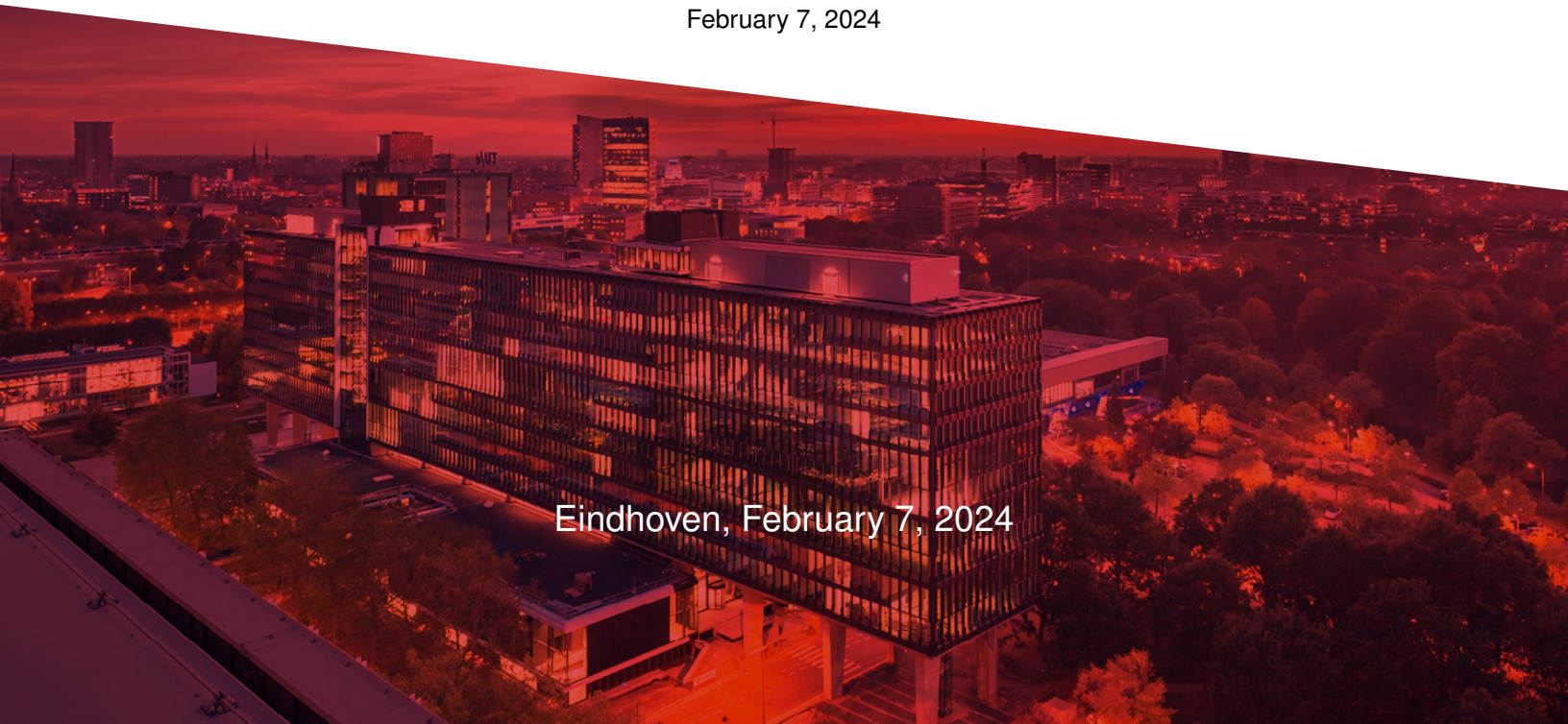
Department of Electrical Engineering
Control System Group

Full Name	Student ID	Master Program
Desen Liang	1736612	Systems and Control

Supervisors:
Prof. dr. ir. Paul M.J.Van den Hof
ir. E.M.M. (Lizan) Kivits

Graduation Committee:
Prof. dr. ir. Paul M.J.Van den Hof
Dr. ir. Maarten Schoukens
Dr. ir. Erik Steur
ir. E.M.M. (Lizan) Kivits

Date of defense:
February 7, 2024



Eindhoven, February 7, 2024

TU/e EINDHOVEN UNIVERSITY OF TECHNOLOGY

Declaration concerning the TU/e Code of Scientific Conduct for the Master's thesis

I have read the TU/e Code of Scientific Conduct[i].

I hereby declare that my Master's thesis has been carried out in accordance with the rules of the TU/e Code of Scientific Conduct

Date

26.01.2024

Name

Desen Liang

ID-number

1736612

Signature

Desen Liang 梁德森

Submit the signed declaration to the student administration of your department.

[i] See: https://www.tue.nl/en/our-university/about-the-university/organization/integrity/scientific-integrity/
The Netherlands Code of Conduct for Scientific Integrity, endorsed by 6 umbrella organizations, including the VSNU, can be found here also. More information about scientific integrity is published on the websites of TU/e and VSNU

CONTENTS

Frequency domain identification in dynamic networks for In-Circuit Testing

Desen Liang
Control system group, Department of Electrical Engineering.
Eindhoven University of Technology
Eindhoven, The Netherlands
d.liang@student.tue.nl

***Abstract*—This thesis introduces a frequency domain identification method for in-circuit testing in Printed Circuit Board Assembly (PCBA), utilizing diffusively coupled dynamic networks. A two-step identification approach is employed: first, estimating the non-parametric frequency response function and noise covariance; second, identifying the parametric network model using either input/output frequency data or the estimated frequency response. A key advantage of this method is its ability to accurately recover all the values of the components within the network. By imposing constraints during the parametric identification phase, the method preserves the structure of the network. This method uses estimated noise covariance to minimize variance, thus obviating the need to select a parametric noise model. Experimental results highlight the method's ability to consistently estimate component values in the complex network with single excitation. The approach is extended to subnetworks, significantly reducing measurement costs in In-Circuit Testing (ICT). Furthermore, a comprehensive ICT procedure is developed that is capable of accurately detecting and diagnosing PCBA defects with minimal excitation. The methodology effectively maintains the relative parameter error of healthy components within $\pm 5\%$. This shows its potential for robust PCBA fault detection and diagnosis.**

***Index Terms*—System identification, Dynamic network, Interconnected systems, Frequency domain, In-Circuit Testing**

I. Introduction

A printed circuit board assembly (PCBA) consists of a board equipped with numerous components that are widely used in various electronic devices, such as computers, smartphones, and televisions. Electronic components are assembled on the board and then soldered. However, various factors can lead to the incorrect assembly of a PCBA. It is vital to ensure that the PCBA is properly assembled before being implemented in complex production. Recently, many pick-and-place machines have been widely used to place components automatically, increasing production efficiency but suffering from the risk of wrong placement. Moreover, problems such as expired components or broken solder can occur even with correct component placement. Consequently, it is imperative to verify the placement and operational condition of each component in PCBA manufacturing.

A. *PCBA testing*

The main content of PCBA testing includes in-circuit testing (ICT), automated optical inspection (AOI) [1] and automated X-ray inspection (AXI) [2]. However, neither AOI nor AXI can detect electrical circuit errors, as they only use images from the camera or X-ray to analyze the PCBA. The common ICT method is to connect the PCBA to external measurement equipment through test points, which are small metal pads on the PCBs used to access the circuit nodes for testing. They are usually connected to the ICT machine probes by spring-loaded pins or needles. The probes can then apply the voltage or current to the test points and measure the circuit response. In this procedure, the ICT method can test the quality of components in PCBA without damaging them. Although ICT looks at the circuit response, it still suffers from inaccurate measurement in densely arranged PCBA due to parasitic effects generated by two closed components with different voltages. In addition, it is difficult to isolate the defect by simply looking at the circuit response, as it can affect the circuit response of other healthy components. Therefore, a more robust and effective test method can be considered to improve ICT for PCBA production.

B. *System identification and dynamic network identification*

The ICT process serves to detect and diagnose faults in PCBA based on the data set applied and collected from the probes. Many different types of methods are used for fault detection and diagnosis (FDD), which can be divided into three parts: data and signal model methods, process model-based methods, and knowledge-based methods [3]. One of the model-based methods is to apply data driven system identification techniques to detect and diagnose faults in dynamic systems. The idea is to identify the value of each component, compare it with the expected value, and analyze the error to verify its correctness. System identification uses statistical methods to build mathematical models of dynamical systems from measured data [4]. Dynamic networks can effectively represent a wide range of dynamic processing models, including complex electrical circuits [5]. Dynamic network identification can be used for electrical circuit testing in PCBA.

Several methods are available to solve this data-driven physical network identification problem based on the time domain. For example, black-box prediction error identification

methods can be used to identify transfer functions from input signals to node signals [6]. State-space models can be used to describe dynamic network systems [7], and identified state-space models can be used to extract the physical values of the components in dynamic networks [8], [9], [10]. However, these methods do not guarantee the accuracy of the estimates due to loss of network structures and do not include consistency analysis. If the estimate can incorporate the network structures, the estimate of interconnection dynamics would be more accurate. Recently, the diffusively coupled network model has been developed to describe dynamic models in which dynamics can be uniquely recovered [11]. A multi-step algorithm was established to identify the full diffusively coupled linear network in structured polynomial models that incorporate the network structure [12] and while they have been extended to the local module (subnetwork) case in [13]. The structured polynomial model approach [12] was applied to improve the ICT process in [14]. This diffusively coupled network model can easily represent the dynamics of electronic components in a complex electrical circuit. However, given that the dynamics of physical components are continuous, it is necessary to explore continuous-time identification methods to identify continuous-time dynamic network models effectively.

C. *Continuous time identification for dynamic network*

Continuous-time (CT) system identification focuses on constructing continuous-time models for dynamical systems using sampled input and output data. There are two main methods in this field: indirect and direct. The indirect approach first establishes a suitable discrete-time (DT) model, which is subsequently converted into a continuous-time model. In contrast, the direct approach derives a continuous-time model directly from the sampled data, avoiding the need for a discrete-time intermediary. Many instrumental variable-based estimation methods have been applied to identify the continuous-time model in both approaches [15] [16]. However, methods for continuous-time dynamic network identification have scarcely been explored. Indirect and direct continuous-time identification approaches for dynamic networks are introduced in [17]. The indirect method introduces errors during the discretization. Conversion can be performed if a transfer function has the same frequency response in both representations [18]. The direct method is an extension of the basic closed-loop instrument variable method (BCLIV) [19] that can be used in continuous-time dynamic networks [20]. Furthermore, a two-step frequency domain identification approach has been developed in [21], which is less complex and scalable for large networks without parameterizing the noise model. However, all of these methods are for the identification of the dynamic network module framework, and there is no guarantee that the value of the interconnected components can be uniquely recovered from the transfer function of the module framework. Thus, a frequency domain identification method was developed for the diffusively coupled continuous-time network in [22]. The frequency domain-based method has many advantages in application: (1) it can avoid approximate derivatives for discretization and directly identify the continuous-time system; (2) it can use the non-parametric noise model directly in estimation instead of parameterizing the noise model; (3) it can directly concatenate multiple experiment data sets; (4) it only requests data samples in the specified frequency range; (5) it can provide straight information of the system dynamics in the frequency response function (FRF).

D. *Research goal*

The research question of this project: *How to estimate the component value accurately with ICT and lower the experiment cost by using dynamic network identification?*

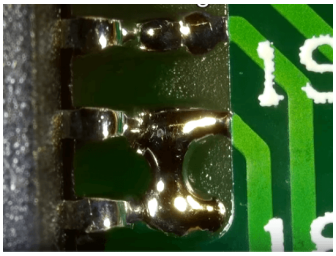
The frequency domain identification method is more suitable to solve the ICT identification problem in PCBA, since it can identify the electronic circuit continuous-time model directly. Straight information on system dynamics can be analyzed from the FRF estimation. Moreover, the noise model is usually unknown for practical applications, so avoiding defining a proper parametric noise model is a significant advantage here. The ability of the diffusively coupled network model identification method to address parasitic effects during ICT has been demonstrated in [14]. This research focuses on the detection and diagnosis of various PCBA defects.

Research goal: To achieve a more robust and effective test method for ICT, the research goal is to develop, implement, and apply a frequency domain identification method for diffusively coupled networks for ICT. This research goal is divided into four tasks as follows.

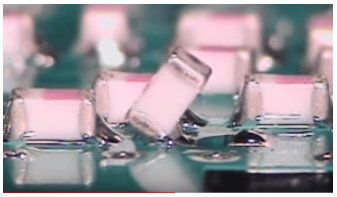
1) In [22], the non-parametric part of applying the frequency domain identification method for diffusively coupled networks has been studied and tested. The first task of this project is to develop the parametric part for identifying full diffusively coupled networks in the frequency domain.
2) To reduce the measurement cost, the second task of this project is to extend the frequency domain identification method for the full network to a part of the network. This part of the network is referred to as a subnetwork in this project.
3) The third task is to implement the whole algorithm in Matlab.
4) The fourth task is to develop a testing procedure for applying the algorithm to the ICT simulation and test it by applying different defects in the simulation.

E. *Common defects in PCBA*

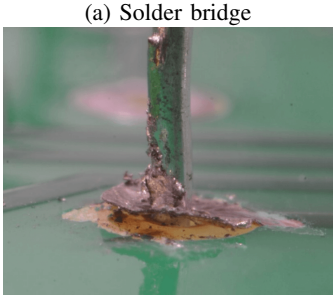
Components in PCBAs are classified into two types: passive and active components [23]. Passive components, including resistors, inductors, capacitors, and transformers, are the most commonly used types in PCBA electrical circuits. Active components comprise transistors and integrated circuits. They may also be implemented in the PCBA to achieve different functions. However, this project focuses on three types of



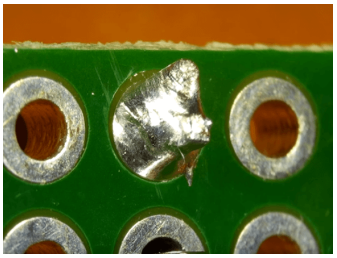
(a) Solder bridge



(b) Tombstoning



(c) Lifted pad



(d) Cold-soldered joint

Figure 1: Examples of common manufacturing defects [24]

electronic passive components: resistors, inductors, and capacitors because they can be easily modeled with the diffusively coupled linear network. Although various defects can arise in PCBA production due to various causes, the most common defects are typically categorized into short circuits, open circuits, and dynamic changes. This project will test the ICT with these three types of defects. Fig. 1 provides examples of these three types of defects.

1) **Short circuit:** Solder bridges, solder balling, etc; Fig. 1a shows an example of a solder bridge. The unwanted solder connections between adjacent pins or pads create a short circuit.
2) **Open circuit:** Missing components, lifted pads, tombstoning, insufficient solder, etc; Fig. 1b and Fig. 1c show the tombstoning and the lifted pad, respectively. The lack of an electrical connection where it should exist leads to an open circuit.
3) **Dynamic changes:** Wrong component placed, skewed component, cold-soldered joints, component cracking, contamination, etc. Dust, oils, or other contaminants in the assembly environment can affect the electrical connection. Fig. 1d shows an example of the cold-soldered joint, which is caused by insufficient heat during soldering, leading to unreliable and weak connections. In this case, although the dynamic of the component itself does not change, the dynamics of the interconnection between the two nodes changes.

F. *Overview of the thesis*

The rest of the thesis consists of 5 parts. Section II introduces the preliminary knowledge as the continuous-time diffusively coupled network model with an electrical circuit network example, the discrete-time network model, and moves the network model to the frequency domain. Section III presents the two-step frequency domain identification methodology with the non-parametric part and the parametric part. The subnetwork identification method in the frequency domain for the diffusively coupled network is shown in Section IV to reduce the measurement cost. The ICT procedure for PCBA with the application of the diffusively coupled network identification algorithm in the frequency domain and the results of various experiments are shown and discussed in Section V. The conclusion and future work discussion are provided in Section VI.

G. *Notation definition*

Some notation are defined that will be used throughout the thesis in this subsection. $A(p)$ is a polynomial matrix with $A(p) = \sum_{\ell=0}^{n_a} A_\ell p^\ell$, it consisting of $n_a + 1$ matrices A_ℓ and (j,k)-th polynomial elements $a_{jk}(p) = \sum_{\ell=0}^{n_a} a_{ik,\ell} p^\ell$. The $a_{ik,\ell}$ is the (j,k)-th element of the matrix A_ℓ. $A_{J\bullet}(p)$ is defined as the polynomial matrix that extracts the first J rows from $A(p)$.

II. Preliminaries

A. *Continuous-time model*

From the last section, the literature indicates that physical networks can often be modeled using undirected graphs with diffusively coupled connections. Here, diffusively coupled connections are interactions that depend only on the difference between node signals, even with a ground or zero node. In particular, these interactions are symmetric in diffusively coupled linear networks, which makes the complex network system easier to analyze. Following the network model in [12], we can describe the full dynamics and topology of diffusively coupled linear networks.

Definition 1. (Diffusively coupled network) The diffusively coupled network consists of L node signals $w(t)$ and K excitation signals $r(t)$. The diffusively coupled network model is described in the following,

$$A(p)w(t) = B(p)r(t) + F(p)e(t), \tag{1}$$

where, p is the differential operator, i.e. $p^\ell w_j(t) = w_j^{(\ell)}(t)$, $w_j^{(\ell)}(t)$ is the ℓ-th order derivative of node signals w_j,

- $A(p) = \sum_{\ell=0}^{n_a} A_\ell p^\ell \in \mathbb{R}^{L\times L}[p]$, with $a_{jk}(p) = a_{kj}(p)$, $\forall k, j$, and $A^{-1}(p)$ is stable.
- $B(p) \in \mathbb{R}^{L\times K}[p]$ is often chosen to be binary and known, which represents each excitation signal directly acting on a distinct node.
- $F(p) \in \mathbb{R}^{L\times L}(p)$ is monic, stable, rational, and the inverse matrix is also stable.

□

Here, the dynamics of known excitation signals $r(t)$ is characterized by the polynomial matrix $B(p)$. Additionally,

unknown disturbance signals are modeled as independent filtered (band-limited) noise. It is assumed that the network is connected, ensuring that there is a path between every pair of nodes. In this context, $F(p)$ is a rational matrix, and $e(t)$ represents a wide-sense stationary white noise process that is independent and identically distributed (iid) at the sampling points with band-limited intersampling behaviors.

Remark 1. The diffusively coupled network model results in a constrained polynomial model, where $A(p)$ is symmetric and non-monic. This captures the diffusively symmetric dynamics in the system. Note that premultiplying (1) with $A^{-1}(p)$ yields the transfer function form model, which is commonly used in the classical system identification [6]. Moreover, this transfer function form model can be transferred to the ARMAX-like or ARX-like model when $F(p)$ is the polynomial or the identity matrix, respectively. This model structure enables the unique extraction of the physical dynamics from the network model. In this project, $F(p)$ is chosen as the polynomial matrix, resulting in a model of ARMAX-like transfer function.

Definition 2. (Continous-time model) For a more simplified representation, the symmetric polynomial matrix $A(p) = X(p) + Y(p)$ is used to represent the dynamic of the diffusively coupled network through which $X(p)$ and $Y(p)$ can be uniquely recovered. We can represent (1) as

$$X(p)w(t) + Y(p)w(t) = B(p)r(t) + F(p)e(t), \qquad (2)$$

where $X(p)$ is a diagonal polynomial matrix with $x_{jj}(p) = \sum_{\ell=0}^{n_x} \mathrm{x}_{jj,\ell} p^{\ell}$; $Y(p)$ is the Laplacian [25] polynomial matrix with $y_{jk}(p) = \sum_{\ell=0}^{n_y} \mathrm{y}_{jk,\ell} p^{\ell}$. n_x and n_y are the order of network dynamics. The terms $\mathrm{x}_{jj,\ell}$ and $\mathrm{y}_{jk,\ell}$ represent the (j,j) diagonal and (j,k) symmetric elements, respectively, of the ℓ-th order matrix $A(p,\ell)$.

□

Utilizing this network model (2) allows for representing complex electrical circuits within PCBA.

Remark 2. (**Intersample behavior for generated data**) In the identification of continuous-time models, the preferred choice is often a band-limited (BL) measurement setup [26]. Here, a BL signal denotes a signal that contains no frequencies higher than a certain limit, known as the Nyquist frequency (half the sampling rate). This implies that the signal can be perfectly reconstructed from its samples if it is band-limited and sampled above the Nyquist rate. All continuous-time model identifications in this project use BL input signals and the BL measurement setup (anti-aliasing filtered). The experimental process in this project involves generating time-domain input data, transforming it to the frequency domain below the Nyquist frequency, simulating the model within this frequency range, and finally transfer the input/output frequency data back to the time domain to establish a continuous-time band-limited data set.

B. A continuous-time example

In this subsection, we present a second-order continuous-time diffusively coupled network model, which is illustrated with a real-world example: the classical resistor-inductor-capacitor (RLC) circuits depicted in Fig. 2.

In Fig. 2, the current flowing through a resistor is determined solely by the voltage difference between the nodes it joins. Diffusively coupling in the context of network theory refers to a type of interconnection between nodes where the coupling or interaction strength is proportional to the difference between the node signals, e.g. the current flow (I_{13}) in R_{13} is proportional to the voltage difference ($U_1 - U_3$) between node 1 and node 3 ($I_{13} = \frac{(U_1 - U_3)}{R_{13}}$). For linear components, this kind of coupling results in symmetric interactions, meaning that the influence of node 1 on node 3 is the same as that of node 3 on node 1.

For the RLC circuits, using Ohm's law, laws of capacitance and inductance can provide the current flowing through each type of component,

$$\begin{aligned} I_R(t) &= \frac{U(t)}{R}, \\ I_C(t) &= C\frac{dU(t)}{dt}, \\ U(t) &= L\frac{dI_L(t)}{dt}, \end{aligned} \qquad (3)$$

where L is the inductance, R is the resistance, $I(t)$ is the current, $U(t)$ is the voltage, and C is the capacitance. According to Kirchhoff's current law (KCL), the total current

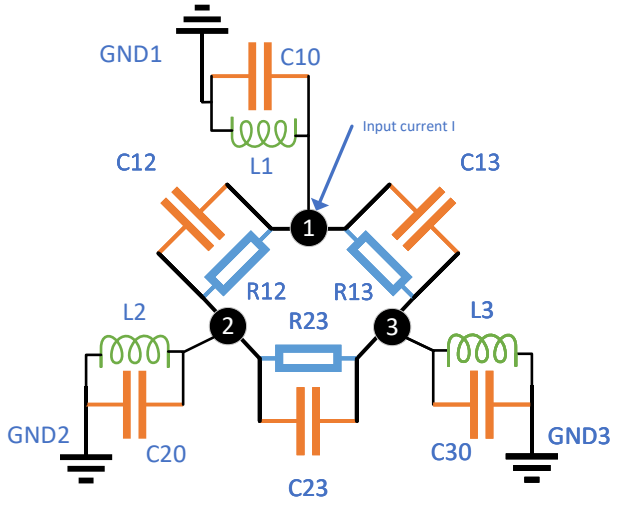


Figure 2: A RLC circuit network with inductors (L_j), resistors (R_{jk}), capacitors (C_{jk}) and ground nodes(GND_j)

entering a given node equals the total current leaving that node. Applying (3) with KCL to the RLC circuit networks, an integro-differential equation can be derived to describe the behavior of this RLC circuit in diffusively coupled network form,

$$\begin{aligned} &C_{j0}\dot{U}_j(t) + \sum_{k\in\mathcal{N}_j} C_{jk}\left(\dot{U}_j(t) - \dot{U}_k(t)\right) + \frac{1}{R_{j0}}U_j(t) \\ &+ \sum_{k\in\mathcal{N}_j} \frac{1}{R_{jk}}\left(U_j(t) - U_k(t)\right) + \frac{1}{L_{j0}}\int_{t_0}^{t} U_j(t)dt \\ &+ \sum_{k\in\mathcal{N}_j} \frac{1}{L_{jk}}\left(\int_{t_0}^{t} U_j(t)dt - \int_{t_0}^{t} U_k(t)dt\right) = I_j(t). \end{aligned} \tag{4}$$

Differentiating this equation with respect to time produces a second-order differential equation devoid of integral terms. Here, the voltage $U(t)$ is considered as the output node signal and the current $I(t)$ is treated as the input signal,

$$\begin{aligned} &C_{j0}\ddot{U}_j(t) + \sum_{k\in\mathcal{N}_j} C_{jk}\left(\ddot{U}_j(t) - \ddot{U}_k(t)\right) + \frac{1}{R_{j0}}\dot{U}_j(t) \\ &+ \sum_{k\in\mathcal{N}_j} \frac{1}{R_{jk}}\left(\dot{U}_j(t) - \dot{U}_k(t)\right) + \frac{1}{L_{j0}}U_j \\ &+ \sum_{k\in\mathcal{N}_j} \frac{1}{L_{jk}}\left(U_j(t) - U_k(t)\right) = \dot{I}_j(t), \end{aligned} \tag{5}$$

where $U_j(t), j = 1, 2, ..., N$ are the N interconnected node signals; $L_{jk} \geq 0, R_{jk} \geq 0, C_{jk} \geq 0, L_{jj} = 0, R_{jj} = 0, C_{jj} = 0$ are the real-valued coefficients of the whole coefficient matrix L, R and C; $L_{jk} = L_{kj}$, $R_{jk} = R_{kj}$ and $C_{jk} = C_{kj}$ are the diffusively coupled constraints; $\mathcal{N}_j$ is the set of indexes of the node signals $U_k(t)$ with connection to $U_j(t)$, $k \neq j$; $I_j(t)$ are the external input signals and $\dot{I}_j(t)$ are the first order derivative of the input signals; $\ddot{U}_j(t)$ and $\dot{U}_j(t)$ are the second order and the first order derivative of the node signals $U_j(t)$ with respect to time t, respectively.

Example 1. To simply illustrate the diffusively coupled structure of this example RLC circuit in Fig. 2, collect the components in matrix form according to (5),

$$\begin{aligned} &A_2 = \\ &\begin{bmatrix} C_{10}+C_{12}+C_{13} & -C_{12} & -C_{13} \\ -C_{12} & C_{20}+C_{12}+C_{23} & -C_{23} \\ -C_{13} & -C_{23} & C_{30}+C_{13}+C_{23} \end{bmatrix}, \\ &A_1 = \begin{bmatrix} \frac{1}{R_{12}}+\frac{1}{R_{13}} & -\frac{1}{R_{12}} & -\frac{1}{R_{13}} \\ -\frac{1}{R_{12}} & \frac{1}{R_{12}}+\frac{1}{R_{23}} & -\frac{1}{R_{23}} \\ -\frac{1}{R_{13}} & -\frac{1}{R_{23}} & \frac{1}{R_{13}}+\frac{1}{R_{23}} \end{bmatrix}, \\ &A_0 = \begin{bmatrix} \frac{1}{L_{10}} & 0 & 0 \\ 0 & \frac{1}{L_{20}} & 0 \\ 0 & 0 & \frac{1}{L_{30}} \end{bmatrix}, \\ &A_2\begin{bmatrix}\ddot{U}_1(t)\\ \ddot{U}_2(t)\\ \ddot{U}_3(t)\end{bmatrix} + A_1\begin{bmatrix}\dot{U}_1(t)\\ \dot{U}_2(t)\\ \dot{U}_3(t)\end{bmatrix} + A_0\begin{bmatrix}U_1(t)\\ U_2(t)\\ U_3(t)\end{bmatrix} = \begin{bmatrix}1\\0\\0\end{bmatrix}\dot{I}(t). \end{aligned} \tag{6}$$

From the above matrix form, it can be clearly observed that $A_i, i = 0, 1, 2$ is a symmetric matrix with respect to diffusively coupled constraints; the sums of the diffusively coupled constraints are shown on the diagonal of the matrices; the components connected to the ground are only shown on the diagonal of the matrices. Consequently, each A_i matrix can be decomposed into two distinct matrices with useful properties: a diagonal matrix X and a Laplacian matrix Y. For example, A_2 can be separated into $X_2 + Y_2$ as

$$\begin{aligned} X_2 &= \begin{bmatrix} C_{10} & 0 & 0 \\ 0 & C_{20} & 0 \\ 0 & 0 & C_{30} \end{bmatrix}, \\ Y_2 &= \begin{bmatrix} C_{12}+C_{13} & -C_{12} & -C_{13} \\ -C_{12} & C_{12}+C_{23} & -C_{23} \\ -C_{13} & -C_{23} & C_{13}+C_{23} \end{bmatrix}, \end{aligned} \tag{7}$$

where, for the (i, j) -th element of X_2 and Y_2, (i, k) -th element of A_2, we have

$$\begin{aligned} X_{2_{ij}} &= \begin{cases} 0, & i \neq j \\ A_{2_{ii}} + \sum_{i\neq k} A_{2_{ik}}, & i = j \end{cases} \\ Y_{2_{ij}} &= \begin{cases} A_{2_{ij}}, & i \neq j \\ -\sum_{i\neq k} A_{2_{ik}}, & i = j. \end{cases} \end{aligned} \tag{8}$$

Given the A_i matrices, this observation allows a straightforward and unique determination of all the values of the components. Subsequently, the second-order RLC circuit model demonstrates that a complicated RLC circuit can be described as a diffusively coupled network. Additionally, this reveals the core idea of the project: by identifying the corresponding network model, we can estimate the dynamics of each individual component in the RLC circuit. Now, having described the diffusively coupled model in the time domain, the subsequent step is to transfer it to the frequency domain. This enables the handling of both discrete-time and continuous-time models through a unified approach.

C. *Discrete-time model*

The discrete-time model is the mapping from the discrete-time input sequence to the output sequence. To reconstruct the continuous-time signals from the discrete-time sequence, digital-to-analog converter (DAC) are been used. Thus, the DT model below is already incorporating the converter. The diffusively coupled discrete-time network model is defined following the same procedure as in [12]. The backward difference method is used to map the continuous-time model to discrete-time

$$\frac{dw(t)}{dt} \approx \frac{w(t_k) - w(t_{k-1})}{T_s}, \tag{9}$$

where with integer k and sampling time T_s, $t_k = kT_s$ are the discrete time instants. For ℓ-order backward difference approximation, the ℓ-th order derivative of node signals $w(t)$ is given as

$$w^{\ell}(t_k) \approx T_s^{-\ell} \sum_{i=0}^{\ell} (-1)^i \binom{\ell}{i} w(t_{k-i}) \text{ with } \binom{\ell}{i} = \frac{\ell!}{i!(\ell-i)!}. \tag{10}$$

Notice that the intersample behavior for identifying the DT model is defined as this first-order approximation for the time derivative in continuous time to discrete time. Therefore, the DT model here is already defined as including a kind of first-order hold linear behavior DAC. The DT data set in this project is generated from this DT model. In this context, the DT data set serves to validate the frequency domain identification outcomes by comparing them with the results obtained using the time domain multi-step (MS) algorithm described in [12].

Definition 3. (Discrete-time model) The continuous-time model in (2) is approximated by using the backward difference method,

$$X(q^{-1})w(t) + Y(q^{-1})w(t) = B(q^{-1})r(t) + F(q^{-1})e(t), \tag{11}$$

where, q^{-1} is the time shift operator, e.g. $q^{-1}w(t_k) = w(t_{k-1})$. Substituting (10) into the (2) and collect the parameters of same differential order, $x_{jj}(q^{-1}) = \sum_{\ell=0}^{n_x} \bar{\mathrm{x}}_{jj,\ell} q^{-\ell}$ and $y_{jk}(q^{-1}) = \sum_{\ell=0}^{n_y} \bar{\mathrm{y}}_{jk,\ell} q^{-\ell}$ in the discrete-time model are given with

$$\begin{aligned} \bar{\mathrm{x}}_{jj,\ell} &= (-1)^{\ell} \sum_{i=\ell}^{n_x} \binom{i}{\ell} T_s^{-i} \mathrm{x}_{jj,i}, \\ \bar{\mathrm{y}}_{jk,\ell} &= (-1)^{\ell} \sum_{i=\ell}^{n_y} \binom{i}{\ell} T_s^{-i} \mathrm{y}_{jk,i}. \end{aligned} \tag{12}$$

□

Remark 3. Note that discretization would not influence the matrix property of $A(p)$, so $A(q^{-1})$ is also the polynomial matrix that can be uniquely divided into a diagonal polynomial matrix and a polynomial Laplacian matrix. The parameters in $B(q^{-1})$ are no longer binary, $B(p)$ and $F(p)$ will discretize into $B(q^{-1})$ and $F(q^{-1})$ following the same procedure as above.

D. *Frequency domain model*

Consider the diffusively coupled network system described in (1). The discrete Fourier transform (DFT) is used to transfer the time-domain model into the frequency-domain model. The DFT $W(k)$,$R(k)$,$E(k)$ of N samples of input-output signals and noise $w(t)$, $r(t)$, $e(t)$ are defined as

$$\begin{aligned} X(k) &= \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} x(t) e^{-j2\pi kt/N}, \\ &\text{with } x = w, r, e \text{ and } X = W, R, E, \end{aligned} \tag{13}$$

where the DFT $E(k)$ for $k = 1, 2, ..., N/2 - 1$ has some important properties as [27]:

1) $E(k)$ is uncorrelated over k, circular complex distributed ($\mathbb{E}\{E^2(k)\}$=0), for any N if $e(t)$ has finite second order moments. This property is sufficient to estimate the noise covariance matrix.
2) $E(k)$ is asymptotically ($N \to \infty$) independent over k, circular complex normally distributed if $e(t)$ exists moments of any order. This property is used to calculate the uncertainty bounds on the measured frequency response function with a given confidence level (see Section 5 in [28]).
3) $E(k)$ is independent over k, circular complex normally distributed for any N if $e(t)$ is Gaussian [28]. This is needed to use the noise covariance matrix as the frequency weighting in the parametric identification of Section III-B .

Definition 4. (Frequency domain model) The model description of the diffusively coupled network in the frequency domain is defined as,

$$A(\Omega_k)W(k) = B(\Omega_k)R(k) + F(\Omega_k)E(k) + C(\Omega_k), \tag{14}$$

if we take the $A^{-1}(\Omega_k)$ of both side of the equation, it gives

$$\begin{aligned} W(k) = &\underbrace{A^{-1}(\Omega_k)B(\Omega_k)}_{G(\Omega_k)} R(k) + \underbrace{A^{-1}(\Omega_k)F(\Omega_k)}_{H(\Omega_k)} E(k) \\ &+ \underbrace{A^{-1}(\Omega_k)C(\Omega_k)}_{T(\Omega_k)}, \end{aligned} \tag{15}$$

where Ω_k is the frequency variable corresponding to frequency index k which is defined as $\Omega_k = j\omega_k$ in continuous time and $\Omega_k = e^{j\omega_k T_s}$ in discrete time with T_s is the sampling time, $\omega_k = 2\pi k f_s/N$ and f_s is the sampling frequency [27], [29]; $G(\Omega_k)$ is the $L \times K$ diffusively coupled network system and $H(\Omega_k)$ is the $L \times L$ noise model; $T(\Omega_k)$ is the transient term as $L \times 1$ vector that include the system and noise transient in time domain which cause leakage errors in measurements of FRF in frequency domain [28].

□

Remark 4. Here, the leakage errors in FRF measurements can be split into two contributions: one is sudden change

(transient) errors in the time domain that are not captured well by the DFT and the other one is the interpolation errors of the FRF over neighboring frequency points in the frequency domain [30].

Henceforth, the diffusively coupled network model has been introduced with the RLC circuit example. This shows the possibility of estimating the values of the RLC components by estimating the corresponding diffusively coupled network model. The common diffusively coupled network in the time domain and the corresponding frequency domain model have also been described. The next step is to develop an identification procedure to estimate the diffusively coupled network in the frequency domain, which is able to identify both the DT model and the CT model.

III. Frequency-domain Identification for diffusively coupled network

This section presents a two-stage identification approach based on the frequency domain to estimate the diffusively coupled network. The FRF and noise covariance are estimated in the first step using a non-parametric methodology. In the second step, this non-parametric noise variance estimation acts as the frequency weighting to estimate the parametric network model. This two-step approach can offer two main advantages: (1) there is no need to parameterize the noise model so that we can avoid choosing a proper model structure for the noise as well as tune the order of the parametric noise model; (2) once the transient term is suppressed in the non-parametric model, the FRF estimation can be used for a curve fitting type procedure with the "transient-free" FRF curve or extracting the "transient-free" input/output data to identify the parametric model.

A. *Non-parametric identification*

As mentioned above, the first step is to find the FRF and the noise covariance estimation. However, when using an arbitrary excitation (e.g., non-periodic signal white noise) to estimate the non-parametric FRF and the noise covariance matrix, the leakage errors of the system and noise will be introduced by transforming the time domain data to frequency domain data via the DFT. In the classical method, different windows are applied to remove the leakage errors. Still, it will also introduce an extra error due to the averaging process for neighboring frequencies [31]. It is hard to trade off between leakage error elimination and noise suppression. Therefore, we have chosen to implement the Local Polynomial Method (LPM), a more advanced approach that accurately estimates the non-parametric FRF and noise covariance matrix, even under arbitrary excitation.

Following the classical LPM method [27], we can use a low-order Taylor expansion to locally approximate smooth frequency functions; see Fig. 3. The FRF $G(\Omega_k)$ and the transient $T_(\Omega_k)$ of the network model in (15) are approximated

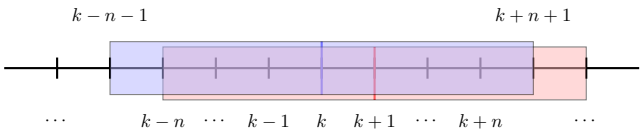


Figure 3: The local polynomial method window at frequency k (blue rectangle) and frequency $k+1$ (red rectangle) in the $2n+1$ DFT frequencies band.

around the selected central frequency k at each frequency band $k+r$ with $r=-n,-n+1,...0,...,n-1,n$,

$$\begin{aligned} W(k) &= \left[G\left(\Omega_k\right)+\sum_{s=1}^{\tau} g_s(k) r^s\right] R(k+r)+V(k+r) \\ &\quad + T\left(\Omega_k\right)+\sum_{s=1}^{\tau} t_s(k) r^s \\ &= \Theta Z(k+r)+V(k+r), \end{aligned} \tag{16}$$

where, τ is the order of the Taylor expansion; g_s and t_s are the polynomial coefficients; $L\times(\tau+1)(K+1)$ matrix Θ collects all the polynomial coefficients and $(\tau+1)(K+1)\times 1$ vector Z_n collects the input data. Collecting (16) in the frequency band for $r=0,\pm1,\pm2,...,\pm n$ ($2n+1$ samples) into a matrix version gives the following:

$$W_n = \Theta Z_n + V_n. \tag{17}$$

The parameter estimation matrix $\hat{\Theta}$ is obtained by solving the least squares cost function as,

$$\hat{\Theta} = \min_{\Theta} \left[W_n - \Theta Z_n\right]^H \left[W_n - \Theta Z_n\right], \tag{18}$$

such that

$$\hat{\Theta} = W_n Z_n^H (Z_n Z_n^H)^{-1}, \tag{19}$$

where, for any complex matrix M, M^H is the Hermitian conjugate transpose of M. Notice that (19) needs to be solved using a numerically stable inversion method, e.g., the singular value decomposition (SVD). Here, the FRF estimation $L\times K$ matrix $\hat{G}(\Omega_k)$ can be extracted from $\hat{\Theta}$ as

$$\hat{G}\left(\Omega_k\right) = \hat{\Theta}_{[:,1:K]}, \tag{20}$$

where $[:,1:K]$ represents the first K columns extracted from the matrix $\hat{\Theta}$.

Finally, substituting the estimation of the polynomial coefficients into (17), we can get the non-parametric noise estimation,

$$\hat{V}_n = W_n - \hat{\Theta} Z_n = W_n[I_{2n+1} - Z_n^H (Z_n Z_n^H)^{-1} Z_n], \tag{21}$$

with this, the estimation of the noise covariance $\hat{C}_V(k)$ is given as,

$$\hat{C}_V(k) = \frac{\hat{V}_n \hat{V}_n^H}{2n+1-(\tau+1)(K+1)}. \tag{22}$$

Remark 5. In this paper, the non-parametric part is implemented using the frequency domain toolbox in [32].

The order τ of the Taylor expansion is chosen as 3 to maintain the balance of bias and variance. The quality of the noise model depends only on the following factors: frequency band width $2n+1$, and order τ of the local polynomial approximation. This can prevent tuning and finding a proper structure for estimating the noise model.

B. Parametric identification

The previous subsection introduced the LPM to estimate the FRF and the noise covariance of the diffusively coupled network model. In this subsection, two criteria can be considered to estimate a parametric network model. The first approach involves utilizing the input/output data $R(k)$ and $W(k)$ to formulate a criterion, incorporating the noise covariance estimation $\hat{C}_V(k)$ as frequency weighting, as detailed in Section III-B2. The alternative method applies non-parametric FRF estimation for parametric modeling, treating the noise covariance estimation $\hat{C}_V(k)$ as frequency weighting, as described in Section III-B3. Before we present the two criteria in the frequency domain, the classical prediction error method in the time domain is introduced first to understand the parametric identification criterion better and compare it with the frequency domain based method.

1) Model set and the prediction error method

The network models that will be identified are collected into a network model set. The model set is defined as a collection of mathematical models that are considered candidates for representing the behavior of the system,

$$\mathcal{M} = \{M(\theta), \theta \in \Psi \subset \mathbb{R}^d\}, \tag{23}$$

$$M(\theta) = (A(p,\theta), B(p,\theta), C(p,\theta), F(p,\theta)), \tag{24}$$

where θ includes all the unknown parameters of the diffusively coupled network model A, B, F, and C in the frequency domain. An equivalent network model is also given but in discrete time,

$$M(\theta_t) = (A(q^{-1},\theta_t), B(q^{-1},\theta_t), F(q^{-1},\theta_t)), \theta_t \in \Psi. \tag{25}$$

Here, 'equivalent' refers to the discrete-time model representing the same mapping as the continuous-time model for the input/output data set. Furthermore, a data generating system $\mathcal{S}$ is defined by th e model,

$$\mathcal{S} = (A^0(p,\theta^0), B^0(p,\theta^0), F^0(p,\theta^0)), \tag{26}$$

if the true parameters $\theta^0 \in \Psi$, then the data generating system is in the model set $\mathcal{S} \in \mathcal{M}$.

From the time domain prediction error method (PEM) [6], we know that parametric identification can be chosen as

$$\hat{\theta}_t = arg\min_{\theta} \frac{1}{N} \sum_{t=0}^{N-1} \varepsilon(t,\theta_t)^T \varepsilon(t,\theta_t), \tag{27}$$

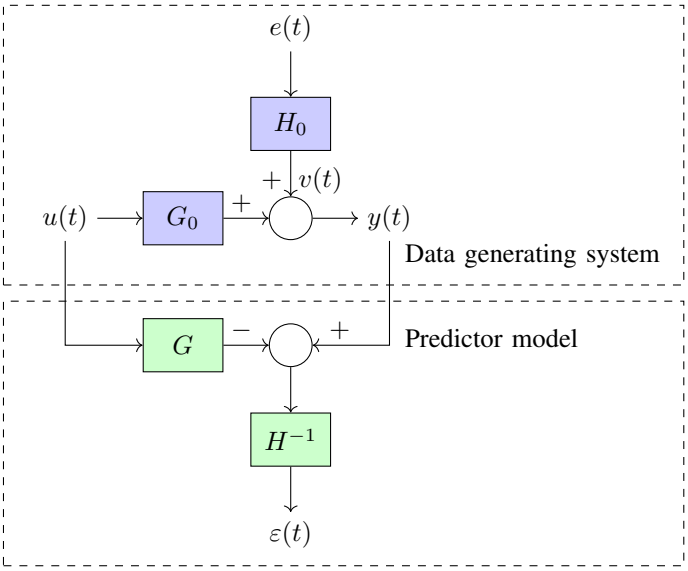


Figure 4: The prediction error framework (the data generating system above, the predictor model below)

where $\varepsilon(t,\theta_t)$ is the one-step-ahead prediction error for the system $y(t) = G(q^{-1})u(t) + H(q^{-1})e(t)$, and $\varepsilon(t,\theta_t)$ is given by,

$$\varepsilon(t,\theta_t) = H(q^{-1},\theta_t)^{-1}\left[y(t) - G(q^{-1},\theta_t)u(t)\right], \tag{28}$$

where G is the plant model with the time-shift operator q and the parameter collection θ_t. The idea is to find the model parameters that minimize the difference between the predicted output of the model and the actual measured output of the real system over a given data set. The structure of the prediction error framework is shown in Fig. 4.

Note that this kind of prediction error framework simultaneously identifies the parametric plant model and the noise model from the data set. It is best to obtain a consistent noise parametric model to improve the consistency of the parametric plant model estimation. The multistep algorithm to parameterize the diffusively coupled network model in [12] is also based on this time domain prediction error method, which also requires the identification of the parametric noise model and $\mathcal{S} \in \mathcal{M}$. However, PEM is more commonly applied to discrete-time models, which suffer from extra discretization errors. Additionally, as network complexity increases, evidenced by high-order dynamics or numerous nodes, determining an optimal structure for the network noise model and fine-tuning its parameters becomes increasingly difficult. Hence, employing a non-parametric noise model that maintains its quality independent of the parametric plant model offers a significant benefit compared to the prediction error framework. Moreover, since the noise model is not parameterized, the noise model is independent of the plant model, and we only need $G_0 \in \mathcal{G}$ for the consistent estimation, with $G_0 = \left\{A^0(p,\theta^0), B^0(p,\theta^0)\right\}$ and $\mathcal{G} = \{A(p,\theta), B(p,\theta), C(p,\theta)\}$ (G_0 and $\mathcal{G}$ are the

subset of $\mathcal{S}$ and $\mathcal{M}$, respectively). Next, two frequency domain identification criteria are discussed for the diffusively coupled network. The time-domain prediction error method is transferred to the frequency domain, which can deal with the continuous-time model without discretization and using the non-parametric noise model instead of the parametric one.

Before we present the two criteria, it is imperative to ensure that specific key conditions are met in preparation for identifying the parametric model.

Condition 1. To satisfy identifiability and consistent estimation, we have the following conditions for the diffusively coupled network according to [12]:

1) The polynomials $A(p)$ and $B(p)$ are left-coprime.
2) There exists at least one diagonal and full-rank matrix of the matrix set $\{A_0, A_1, ..., A_{n_a}, B_0, B_1, ..., B_{n_b}\}$.
3) There exists at least one excitation signal, $K \geq 1$.
4) There exists at least one constraint on the parameters of $A(p,\theta)$ and $B(p,\theta)$ that ensures that $\Gamma\theta = \gamma \neq 0$.
5) The system is excited at all frequencies of interest.

Remark 6. The first condition ensures that there are no common factors between $A(p,\theta)$ and $B(p,\theta)$. The non-monic of the polynomial matrix $A(p,\theta)$ is addressed in the second and fourth conditions, which require the inclusion of at least one identity matrix in the polynomial, modified by a constant scaling factor. The second condition has been relaxed so that there exists a permutation matrix P that leads to $[A_0\ A_1\ ...\ A_{n_a}\ B_0\ B_1\ ...\ B_{n_b}]\,P = [D\ U]$ with D square, diagonal, and full rank in [33]. Normalizing the model with this specific scaling factor guarantees the uniqueness of this model representation. Consequently, the fourth condition is crucial to find a unique representation of this model during the parametric identification procedure. A more detailed explanation of optimization with constraints will be provided in Section III-B4.

2) Identification criterion for f-domain input/output data

One of the identification criteria is to use the frequency domain data set $W(k)$ and $R(k)$, which is similar to the prediction error method using the input/output data set in the time domain. According to Parseval's Theorem, the prediction error in the frequency domain with DFT is given based on (15), (27), and (28),

$$\varepsilon(k,\theta) = H(\Omega_k,\theta)^{-1}\left[W(k) - G(\Omega_k,\theta)R(k) - T(\Omega_k,\theta)\right], \tag{29}$$

where k is the frequency point index, $T(\Omega_k,\theta)$ is the transient includes the plant and noise transients. The parametric estimation for the prediction error in the frequency domain is given by,

$$\begin{aligned}\hat{\theta} &= arg\min_{\theta}\frac{1}{N}\sum_{k=1}^{N}||\varepsilon(k,\theta)^H\varepsilon(k,\theta)||_F^2,\\ \varepsilon(k,\theta) &= H(\Omega_k,\theta)^{-1}\left[W(k) - G(\Omega_k,\theta)R(k) - T(\Omega_k,\theta)\right],\end{aligned} \tag{30}$$

where, for any complex matrix M, $||M||_F$ is the Frobenius Norm of M.

Remark 7. With M_{ij} are the elements of M, and m, n are the dimensions of M, the definition of Frobenius Norm is given in [34] as

$$||M||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|M_{ij}|^2}, \tag{31}$$

where, $|M_{ij}|^2$ denotes the amplitude square of the complex number M_{ij}.

Nevertheless, for the unknown noise model, this ARMAX-like nonlinear-in-parameters structure leads to a nonconvex optimization problem which might suffer from local minima, and the result strongly depends on the initial conditions. To deal with this computational problem, we can premultiply $A(\Omega_k)$ in each step to make it linear in the parameters and then divide $A(\Omega_k)^{i-1}$ from the last step as a frequency weighting. This method is well known as the SK-algorithm [35]. Moreover, since $|H(\Omega_k,\theta)|^2$ only plays a role as frequency weighting here, the noise covariance estimation $\hat{C}_V(k)$ from the non-parametric estimation can be used to avoid parameterizing the noise model, but reaches the same goal. Now, if the noise signals $E(k)$ satisfy the conditions given in Section. II-D, the noise can be modeled as $V(k) = H(\Omega_k,\theta)E(k)$, then we have the following proposition,

$$\mathbb{E}\{E(k)\} = 0, \quad \mathbb{E}\{E^2(k)\} = 0, \quad \mathbb{E}\{|E(k)|^2\} = \sigma_E^2, \tag{32}$$

such that

$$\begin{aligned}&\mathbb{E}\{V(k)\} = 0, \quad \mathbb{E}\{V^2(k)\} = 0,\\ &\mathbb{E}\{|V(k)|^2\} = |H(\Omega_k,\theta)|^2\,\sigma_E^2 = \sigma_V^2.\end{aligned} \tag{33}$$

Since the variance σ_E^2 is just a constant number, it can be considered scaling the weightings for the entire frequency range. If we ignore this term of the cost function, it will not influence the converged result. Therefore, the criterion can be given by replacing $|H(\Omega_k,\theta)|^2$ with noise covariance estimation $\hat{C}_V$,

$$\begin{aligned}\hat{\theta} &= arg\min_{\theta}\frac{1}{N}\sum_{k=1}^{N}||M_1((k),\theta)||_F^2,\\ M_1(k,\theta) &= \underbrace{\left[\hat{C}_V(k)^{\frac{1}{2}}A(\Omega_k,\theta)^{i-1}\right]^{-1}}_{\mathfrak{W}}\\ &\quad\times\left[A(\Omega_k,\theta)W(k) - B(\Omega_k,\theta)R(k) - C(\Omega_k,\theta)\right].\end{aligned} \tag{34}$$

At first glance, this criterion might work well. However, the parametric result of this approach is not consistent for the

normally distributed noise signal $v(t)$, as $W(k)$ and $\hat{C}_V(k)$, although uncorrelated, are not independently distributed. Since the FRFs are approximated locally in overlapping finite frequency bands, the transient error and the noise covariance matrix estimation are correlated over the frequency.

To address this problem, we employ a method that involves the generalized sample covariance and sample mean. These are derived from the LPM estimation of the DFT of node signals $W(k)$ and its associated covariance, given their asymptotic independence [27]. Utilizing the non-parametric frequency response estimation matrix $\hat{\Theta}$ in (19), the sample mean $\hat{W}(k)$ of $W(k)$ is calculated as

$$\hat{W}(k) = \hat{G}\left(\Omega_k\right) R(k) + \hat{T}\left(\Omega_k\right), \tag{35}$$

where $\hat{G}\left(\Omega_k\right)$ is the FRF estimations obtained from $\hat{\Theta}$ in the non-parametric identification part (20) and $\hat{T}(\Omega_k)$ is the transient term given as $\hat{T}\left(\Omega_k\right) = \hat{\Theta}_{[:,K(R+1)+1]}$ ($[:, 1:K]$ represents the first K columns extracted from the matrix $\hat{\Theta}$). The sample covariance $\hat{C}_W(k)$ of $\hat{W}(k)$ can be calculated from the noise covariance model obtained from the non-parametric part as

$$\hat{C}_W(k) = (Z_n^H (Z_n Z_n^H)^{-1} Z_n)_{[n+1,n+1]} \hat{C}_V(k). \tag{36}$$

Remark 8. The asymptotical behaviors and the proof details are shown in [36] and Chapter 12 in [27].

By substituting $W(k)$ and $\hat{C}_W(k)$ with $\hat{W}(k)$ and $\hat{C}_W(k)$ in $M_1(k,\theta)$ as presented in (34), we can formulate the concluding criterion for the frequency domain input/output data as follows,

$$\begin{aligned}
&\hat{\theta} = arg\min_{\theta} \frac{1}{N} \sum_{k=1}^{N} ||M_1((k),\theta)||_F^2 , \\
&M_1(k,\theta) = \mathfrak{W}\left[A(\Omega_k,\theta)\hat{W}(k) - B(\Omega_k,\theta)R(k) - C(\Omega_k,\theta)\right], \\
&\text{with } \mathfrak{W} = \left[\hat{C}_W(k)^{\frac{1}{2}} A(\Omega_k,\theta)^{i-1}\right]^{-1}.
\end{aligned} \tag{37}$$

Collecting all the input/output data for all frequencies with the network structure, the weighting $\mathfrak{W}$ in the regression matrix Q_1 and parameters in vector θ_1 for each iteration,

$$M_1 = Q_1 \theta_1, \tag{38}$$

where, θ_1 includes all the model parameters in $A(\Omega_k,\theta)$, $B(\Omega_k,\theta)$ and $C(\Omega_k,\theta)$. The parameters are estimated by solving a least-square optimization as

$$\hat{\theta} = \min_{\theta_1} \theta_1^H Q_1^H Q_1 \theta_1 \tag{39}$$

Remark 9. The structures of the regression matrix Q_1 and the parameter vector θ_1 are given in the Appendix. VII-A.

*3) **Identification criterion for FRF data***

As the FRF and the noise covariance estimation have been obtained from the non-parametric part by LPM, these data can be used as prior knowledge to identify the parametric model, for the noise and the transient term have been removed from the FRF. A natural idea would be to use the FRF estimation $\hat{G}\left(\Omega_k\right)$ as the target curve [37] that we want to fit and consider the parameterized procedure as a curve-fitting problem with specific constraints on the network structure. The FRF estimation data $\hat{G}(k)$ consist of complex matrices $L \times K$ $\hat{G}_{ij}(k)$ with $i = 1, ..., L$ and $j = 1, ..., K$, and the objective is to deduce the most accurate diffusively coupled network parametric model that complies optimally with the FRF data.

Following Definition 4, the error $M_2(\Omega_k,\theta)$ between the sampled output without transient and the parametric model output is aimed to minimize,

$$\begin{aligned}
M_2(k,\theta) &= H(\Omega_k,\theta)^{-1} \\
&\times \left[\hat{G}(\Omega_k)R(k) - \underbrace{A^{-1}(\Omega_k,\theta)B(\Omega_k,\theta)}_{G(\Omega_k,\theta)} R(k)\right],
\end{aligned} \tag{40}$$

if we directly construct a quadratic optimization problem by taking the square of M_2 and solving it, the result might not reach the global minima for this ARMAX-like nonlinear-in-parameters structure. To avoid constructing the non-convex problem and chase consistency, the same idea as in the first criterion is considered. Pre-multiply $A(\Omega_k,\theta)$ with (40) to make the structure linear in parameters and also multiply the inverse of the last iteration $A(\Omega_k,\theta)^{i-1}$ as one of the frequency weighting. Recall the sample covariance $\hat{C}_W(k)$ in (36), and consider it also as a frequency weighting; the criterion can be given as

$$\begin{aligned}
&\hat{\theta} = arg\min_{\theta} \frac{1}{N} \sum_{k=1}^{N} ||M_2((k),\theta)||_F^2 , \\
&M_2(k,\theta) = \\
&\underbrace{\left[\hat{C}_W(k)^{\frac{1}{2}} A(\Omega_k,\theta)^{i-1}\right]^{-1}}_{\mathfrak{W}} \left[A(\Omega_k,\theta)\hat{G}(\Omega_k) - B(\Omega_k,\theta)\right] R(k).
\end{aligned} \tag{41}$$

Remark 10. For this criterion, we have assumed that the transient term has been estimated as the non-parametric frequency response function collected in $\hat{\Theta}$ of (19). Therefore, the transient term can be ignored when the model is parameterized from the non-parametric FRF estimation $\hat{G}$.

Collecting all the FRF, input data, and the weighting data $\mathfrak{W}$ for all frequencies into the regression matrix Q_2 and parameters in vector θ_2 for each iteration,

$$M_2 = Q_2 \theta_2, \tag{42}$$

where, θ_2 includes all the model parameters in $A(\Omega_k, \theta)$ and $B(\Omega_k, \theta)$. The parameters are estimated by solving a least-square optimization as

$$\hat{\theta} = \min_{\theta_2} \theta_2^H Q_2^H Q_2 \theta_2, \tag{43}$$

Remark 11. The structures of regression matrix Q_2 and parameter vector θ_2 are given in the Appendix. VII-B.

In conclusion, both criteria follows the idea from SK-algorithm to premultiply the $|A(\Omega_k, \theta)|$ with the cost function to make a linear-in-parameters model structure that can find an analytical solution. Then multiply $|A(\Omega_k, \theta)^{i-1}|$ from the last iteration step as frequency weighting, which will not change the convexity of this optimization problem. Therefore, the solution of parametric identification for the diffusively coupled network model is easier to find.

*4) **Lagrangian multiplier optimization***

Since the non-monic property A polynomial matrice and the parameters in B might be partially known, we cannot construct a standard least-squares optimization problem. Two constraints are considered to be added to maintain the symmetric structure of the A matrices and the binary known B matrices and to achieve the identifiable condition for the model and the uniqueness of the solution, which are given in Condition 1. Based on the criterion (37) and (41), the idea of estimating the parametric model is to find the parameters that minimize the cost function with the corresponding criteria, then the parametric estimation is given by

$$\begin{aligned} \hat{\vartheta} = \min_{\vartheta} \vartheta^H Q^H(\phi) Q(\phi) \vartheta, \\ \text{subject to } \Gamma \vartheta = \upsilon, \end{aligned} \tag{44}$$

where, $Q(\phi)$ and ϑ can be either Q_1 or Q_2 with the corresponded θ_1 or θ_2, Γ is the matrix that selecting the parameters and υ is the constraint vector.

Remark 12. The structure of the selecting matrix Γ and the constraint vector υ are shown in the Appendix. VII-C.

Then the Lagrangian multiplier and Karush–Kuhn–Tucker (KKT) conditions [38] can be used to solve this optimization problem, which gives

$$\mathcal{L} = \vartheta^H Q^H(\phi) Q(\phi) \vartheta + \lambda(\Gamma \vartheta - \upsilon), \tag{45}$$

with the conditions,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \vartheta} &= 2Q(\phi)^H Q(\phi) \vartheta + \Gamma^T \lambda = 0, \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= \Gamma \vartheta - \upsilon = 0. \end{aligned} \tag{46}$$

Collecting them into the matrix representation,

$$\begin{bmatrix} 2Q(\phi)^H Q(\phi) & \Gamma^T \\ \Gamma & \mathrm{O}_{d_{13} \times d_{14}} \end{bmatrix} \begin{bmatrix} \vartheta \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathrm{O}_{d_{15} \times 1} \\ \upsilon \end{bmatrix}, \tag{47}$$

so that the estimation of the parameters ϑ is given as,

$$\begin{bmatrix} \hat{\vartheta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} 2Q(\phi)^H Q(\phi) & \Gamma^T \\ \Gamma & \mathrm{O}_{d_{13} \times d_{14}} \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{O}_{d_{15} \times 1} \\ \upsilon \end{bmatrix}, \tag{48}$$

where O are the zero matrices with the dimension $d_{13} \times d_{14} = [\frac{L(L-1)}{2}(n_a+1)+LK(n_b+1)] \times [\frac{L(L-1)}{2}(n_a+1)+LK(n_b+1)]$ and $d_{15} \times 1 = [L^2(n_b+1) + LK(n_b+1) + L(n_c+1)] \times 1$ with $n_c = -1$ when the transient term is ignored for the second criterion in Section. III-B3.

*5) **Computational method***

So far, we already have the basic ideas and two optimization criteria for identifying the parametric diffusively coupled network model in the frequency domain. The objective in the parametric phase is to achieve a unique, real parametric representation of the model despite the complex values of both input/output and FRF data. The challenge at this stage is to determine how we can utilize these complex-valued data to identify the appropriate parametric model.

To find a real value minimizer $\hat{\vartheta}$, we can treat the real and imaginary parts of the data as separate entities but optimize them simultaneously. First, since $Q(\phi)$ collects all data and weights for all frequency samples in three dimensions $L \times [L^2(n_a+1) + LK(n_b+1) + L(n_c+1)] \times F$ ($n_c = -1$, if the transient term is ignored), we can define the number of parameters in the model as $n_\vartheta = L^2(n_a+1) + LK(n_b+1) + L(n_c+1)$. The third-order regression tensor is reshaped as a matrix in $Ln_\vartheta \times F$. Extract the real part and imaginary parts from the regression matrices $Q(\phi)$ separately, the regression matrix $Q_{re}(\phi)$ is defined as

$$Q_{re} = \begin{bmatrix} \mathrm{Re}(Q(\phi)) \\ \mathrm{Im}(Q(\phi)) \end{bmatrix}, \tag{49}$$

where, $\mathrm{Re}()$ extract all the real part value and $\mathrm{Im}()$ extract all the imaginary part value, then Q_{re} stacks these parts on top of each other. Substituted this matrix with real values in (70) gives

$$\begin{bmatrix} \hat{\vartheta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} 2Q_{re}(\phi)^T Q_{re}(\phi) & \Gamma^T \\ \Gamma & \mathrm{O}_{d_{13} \times d_{14}} \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{O}_{d_{15} \times 1} \\ \upsilon \end{bmatrix}. \tag{50}$$

Defined $\mathcal{Q}$ and $\mathcal{V}$ as,

$$\mathcal{Q} = \begin{bmatrix} 2Q_{re}(\phi)^T Q_{re}(\phi) & \Gamma^T \\ \Gamma & \mathrm{O}_{d_{13} \times d_{14}} \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \mathrm{O}_{d_{15} \times 1} \\ \upsilon \end{bmatrix}. \tag{51}$$

In this way, all terms now have real values, so we can focus on solving this Lagrangian optimization problem. However, it is important to note that during implementation, the inversion in (50) tends to be unstable, largely due to two main factors. One reason is the enlargement of the matrix $\mathcal{Q}$ due to increased constraints, but the upper left part $2Q_{re}(\phi)^T Q_{re}(\phi)$ that contains the data and the weighting of the last iteration $A(\vartheta)^{i-1}$ is independent of the constraints. The other elements, filled with 1 or 0, are for selecting constraint parameters and to shape $\mathcal{Q}$ into a

square matrix. Another reason involves the actual $A^0(\theta^0)$ network model. When this model encompasses extremely high or low parameters or is of higher order, it creates a marked discrepancy in values between the upper left part and the rest, thus increasing the condition number of the matrix $\mathcal{Q}$.

Notice that since the discrete-time model can be considered a scaled version compared to the corresponding continuous-time model, the parameter values across various orders converge to a similar level, resulting in enhanced numerical stability compared to the direct use of the continuous-time model. Therefore, the best way to minimize the condition number of matrix $\mathcal{Q}$ for the continuous-time model is to find a scale factor that depends on the specific model. A more effective general approach is to employ the median of the angular frequencies within the relevant frequency band: $\omega_{scale} = \text{median}\{\omega_1, \omega_2, ..., \omega_F\}$ [39]. For example, the second-order matrix $A_2(\vartheta)(j\omega_k)^2$ could be scaled, giving the result $A_2(\vartheta)\omega_{scale}^2(j\omega_k/\omega_{scale})^2$. Therefore, the estimation parameters become $A_2(\vartheta)\omega_{scale}^2$ and the parameters are scaled to improve the numerical stability.

To further improve the stability of the calculation, we use the SVD method to calculate the inverse of this matrix. The singular value decomposition of this matrix is given as,

$$\mathcal{Q} = U_{\mathcal{Q}}\Sigma_{\mathcal{Q}}V_{\mathcal{Q}}^T, \tag{52}$$

such that the numerically stable parametric estimation is given by

$$\hat{\vartheta} = \left[U_{\mathcal{Q}}\Sigma_{\mathcal{Q}}^{-1}V_{\mathcal{Q}}^T\mathcal{V}\right]_{[1:n_\vartheta,1]}. \tag{53}$$

Finally, we also scale each column of the matrix $\mathcal{Q}$ by its 2-norm as mentioned in [27], which can benefit numerical stability. In conclusion, several methods are used to increase the numerical stability during the implementation of optimization.

C. *Codes checking*

Since the key objective is to verify whether the parameters of the PCBA components are correct, it is crucial to ensure the correct implementation of the parametric part. Therefore, we develop some double-checking codes to verify the implementation. For example, find two different ways to construct the key matrices in the optimization problem and then compare the difference between these two-way generated matrices. In Appendix III-B4, we present a way to construct the constraint selection matrix; another way to generate the constraint selection matrix for the symmetric $A(\Omega_k, \vartheta)$ is introduced in Appendix VII-D.

D. *Full network identification Flow*

To conclude the frequency domain full network identification procedure, the following flow is given:

1) **Non-parametric Identification**
   - **Input:** The input and output signals data set, frequency range.
   - **Process:** Using LPM method to estimate $\hat{\Theta}$ in (19) and noise covariance $\hat{C}_V$ in (22).
   - **Output:** The estimated FRF $\hat{G}$ from $\hat{\Theta}$ and the noise covariance $\hat{C}_V$.
2) **Parametric Identification**
   - **Input:** The user-defined parametric model order n_a of $A(p,\theta)$, n_b of $B(p,\theta)$, n_c of $C(p,\theta)$; at least one constraint on non zero parameter e.g., $b_{ij,\ell} = 1$; input/output frequency data; $\hat{G}$ and $\hat{C}_V$.
   - **Process:** If $n_c > -1$, using the input/output data criterion shown in Section III-B2 optimize the weighted least squares following (37) with constraint (44) to estimate parameters $\hat{\vartheta}$, otherwise using the FRF data criterion shown in Section III-B3 the optimization of the weighted least squares following (41) with constraint (44) to estimate parameters $\hat{\vartheta}$.
   - **Output:** Estimation of the full network $\hat{A}(\hat{\vartheta})$ and $\hat{B}(\hat{\vartheta})$.

Remark 13. Notice that this identification flow is fully in the frequency domain to identify the frequency-domain model. Therefore, this flow works the same for the identification of the CT model and the DT model. The differential equations that describe the model have already been converted to algebraic equations by substituting the related frequency variable, as shown in Definition 4.

E. *Simulation results for full network identification*

Several experiments were done in Matlab simulation to test the frequency domain identification for the diffusively coupled full network algorithm. Consider a seven-node RLC network as shown in Fig. 5, where each node has the measurement components connected to the ground node (shown in Table I). The parameters of each component of the interconnections are given in Table II. The orders of the parametric model are $n_a = 2$ and $n_b = 1$, the same as the Example 1. In addition, the parametric transient term is also estimated with the order $n_c = 0$, criterion based on the input/output data shown in Section III-B2 is used here. The CT and DT models share the same order settings. Constraints are given as the known input matrix $B(p)$, the known unconnected parts of the network represented as zero terms A_{zeros} in the matrix $A(p)$ for the CT model ($B(q^{-1})$, $A(q^{-1})$ for the DT model).

Table I: Seven-node Measurement Component Values

Identifier	Value	Unit
C_{j0}	2	μF
R_{j0}	500	Ω
L_{j0}	18	mH

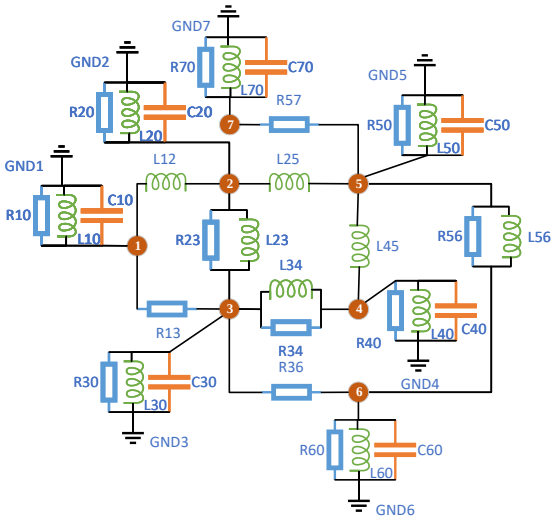


Figure 5: A seven-node RLC network Example with inductors (L_{jk}), resistors (R_{jk}), capacitors (C_{jk}) and ground nodes(GND_j) [14]

Table II: Seven-node Network Component Values

Resistor	Value	Unit	Inductor	Value	Unit
R_{13}	100	Ω	L_{12}	5	mH
R_{23}	200	Ω	L_{23}	10	mH
R_{34}	150	Ω	L_{25}	15	mH
R_{36}	180	Ω	L_{34}	12	mH
R_{56}	160	Ω	L_{45}	20	mH
R_{57}	120	Ω	L_{56}	13	mH

Example 2. Giving the excitation signal $r(t)$ as the independent zero mean white noise with variance $\sigma_r^2 = 1$ entering only node 6, and the normal distributed zero-mean white noise as the noise signal $e(t)$ with variance $\sigma_e^2 = 1$ entering all nodes. The sampling frequency is set at 20000 Hz to cover all the dynamics of the components, and the identification frequency band is set between $f_{min} = 500$ Hz and $f_{max} = 6000$ Hz to promote estimation performance. Given the situation that can only excite one node and measure all nodes, this 7-node RLC circuit can be expressed as a continuous-time diffusively coupled network model following the same procedure as in Example 1, and this CT model is discretized following Definition 3 to obtain the DT model. The CT and DT data sets are generated from the CT and DT models, respectively. To show that the parameters can be consistently identified with a single excitation, we generated a set of experiments with different lengths of the data N. The choice of N and the corresponding set are shown in Table III. Each set of experiments includes 50 Monte Carlo (MC) runs with independent excitation and noise signals.

The box plots of the relative mean squared error (RMSE) of the parameters of the components for each experimental set

Table III: Different data length with experiment set

Set	1	2	3	4	5
N	10^3	2×10^3	4×10^3	8×10^3	16×10^3
Set	6	7	8	9	10
N	32×10^3	64×10^3	128×10^3	256×10^3	512×10^3

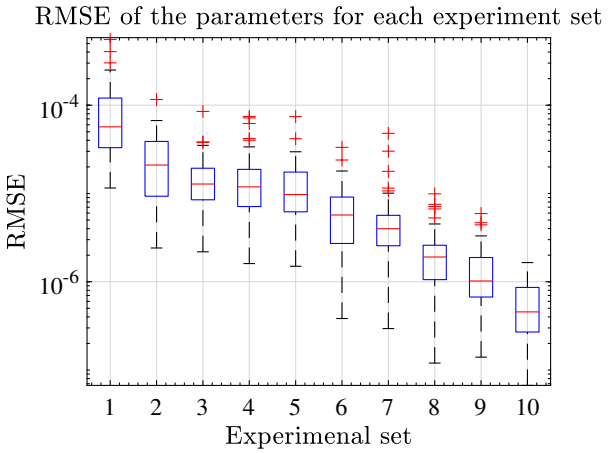


Figure 6: Boxplot of the RMSE of the parameters of the components for each experimental set (CT model)

are shown in Fig. 6 and Fig. 7 with the continuous-time model and the discrete-time model, respectively. The relative mean squared error of the parameters of the components is given as

$$\text{RMSE} = \frac{\|\hat{\theta}_{comp} - \theta^0_{comp}\|_2^2}{\|\theta^0_{comp}\|_2^2}, \tag{54}$$

where $\hat{\theta}_{comp}$ collects the estimated parameters of components and θ^0_{comp} collects the actual parameters of components. From these figures, it can be seen that RMSE decreases as N increases, which supports the achievement of consistent identification. When the data length N extends to infinity,

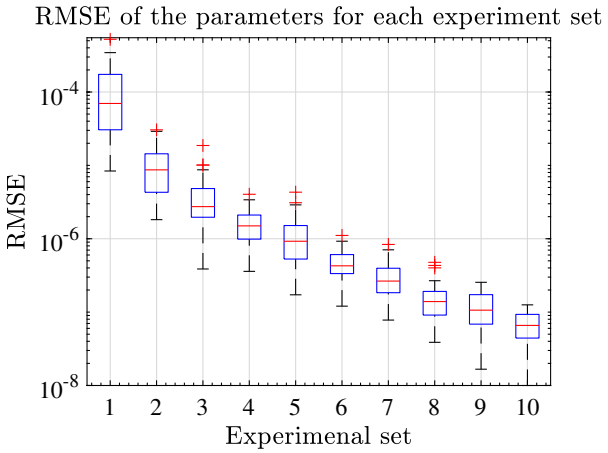


Figure 7: Boxplot of the RMSE of the parameters of the components for each experimental set (DT model)

the estimated parameters converge to the actual parameters. In conclusion, the first task of this project is achieved, a consistent frequency domain identification method is developed for the diffusively coupled network, and the consistency of it is shown by the experiments.

IV. Subnetwork identification for the diffusively coupled network in frequency domain

As detailed in the previous chapter, the frequency domain identification algorithm targets the identification of all interconnection dynamics in diffusively coupled networks, demanding full measurement or excitation data. In the context of ICT procedures, ensuring that two test points are in the PCBA for each component incurs significant measurement costs. Such test points occupy space on the PCBA, leading to increased production costs, complexity in design, and a larger and more cumbersome PCBA. As a result, more advanced ICT systems that use specialized testing methods enable component testing within subnetworks or groups. Testing components in subnetworks is particularly efficient for PCBAs with a high density of components. This approach reduces testing time and can detect certain types of fault more effectively. Currently, testing the subnetworks in the PCBAs can involve a combination of time-domain and frequency-domain response analyses, direct measurement of the component values, etc. However, direct measurement of component values in subnetworks typically requires isolating the component from the rest of the circuit as much as possible. Thus, integrating the advantages of response analyses and parameter estimation without isolating subnetwork components, this chapter presents a frequency domain subnetwork identification algorithm for the diffusively coupled network. This methodology builds on the full network frequency domain identification algorithm for diffusively coupled networks, as outlined in the previous chapter, adapting the time domain subnetwork identification algorithm from [13] to the frequency domain to reduce errors in the estimation of continuous-time RLC circuits. In this section, (1) the immersed network will be introduced to model the network with partial measurement in Section IV-A; (2) the subnetwork identification procedure in the frequency domain will be presented in Section IV-B; (3) this procedure will be concluded in an algorithm flow in Section IV-C; (4) some simulation experiments will be shown to test the algorithm in Section IV-D.

A. *The immersed network*

Identifying the subnetwork with partial measurement can save much measurement cost. In contrast to full network identification, where all nodes are measured, subnetwork identification requires measuring only a specific subset of node signals. Those unmeasured node signals can be removed from the model by using Gaussian elimination. This Gaussian elimination process is known in scholarly texts as Kron reduction [40] or immersion [41] [13]. Several algorithms have been developed to identify the local module or subnetwork of a dynamic network, but without taking into account the structure of the undirected network model [21] [41]. Here, following the literature [13], the immersed network of a diffusively coupled network can be given as follows.

Definition 5. (Immersed diffusively coupled network) Consider a diffusively coupled network model as defined in Definition 1, the node signals $w(t)$ are divided into two groups; one is the measurement groups $w_{\mathfrak{M}}(t)$ with the measurement set $\mathfrak{M} = \{j | w_j \in w_{\mathfrak{M}}\}$, the other is the immersion groups $w_{\mathfrak{I}}(t)$ with the immersion set $\mathfrak{I} = \{j | w_j \notin w_{\mathfrak{I}}\}$. The immersed diffusively coupled network is defined as

$$A_{im}(p) w_{\mathfrak{M}}(t) = B_{im}(p) r(t) + F_{im}(p) e_{\mathfrak{M}}(t), \tag{55}$$

where $A_{im}(p)$ and $B_{im}(p)$ are still polynomial matrices. Here, since we are using the frequency domain identification method, there is no need to parameterize the noise model. We can consider $F_{im}(p)e_{\mathfrak{M}}(t)$ as $V_{im}(t)$.

□

For the two separate groups of node signals, the diffusively coupled network defined in Definition 1 can be divided into the following structure:

$$\underbrace{\begin{bmatrix} A_{\mathfrak{M}\mathfrak{M}} & A_{\mathfrak{M}\mathfrak{I}} \\ A_{\mathfrak{I}\mathfrak{M}} & A_{\mathfrak{I}\mathfrak{I}} \end{bmatrix}}_{A(p)} \underbrace{\begin{bmatrix} w_{\mathfrak{M}}(t) \\ w_{\mathfrak{I}}(t) \end{bmatrix}}_{w(t)} = \underbrace{\begin{bmatrix} B_{\mathfrak{M}} \\ B_{\mathfrak{I}} \end{bmatrix}}_{B(p)} r(t) + \underbrace{\begin{bmatrix} V_{\mathfrak{M}}(t) \\ V_{\mathfrak{I}}(t) \end{bmatrix}}_{F(p)e(t)}, \tag{56}$$

to immersed the unmeasured signals, Gaussian elimination is employed. The immersed signals are expressed as

$$w_{\mathfrak{I}}(t) = A_{\mathfrak{I}\mathfrak{I}}^{-1} \left(B_{\mathfrak{I}} r(t) + V_{\mathfrak{I}}(t) - A_{\mathfrak{I}\mathfrak{M}} w_{\mathfrak{M}}(t) \right), \tag{57}$$

and then substituted with

$$A_{\mathfrak{M}\mathfrak{M}} w_{\mathfrak{M}}(t) + A_{\mathfrak{M}\mathfrak{I}} w_{\mathfrak{I}}(t) = B_{\mathfrak{M}} r(t) + V_{\mathfrak{M}}(t), \tag{58}$$

the immersed model can be obtained. For calculating the inverse of a polynomial matrix, we know that for an arbitrary polynomial matrix $M(p)$, its inverse is given by

$$M^{-1}(p) = \frac{\mathrm{adj}(M(p))}{\det(M(p))}, \tag{59}$$

where $\mathrm{adj}(M(p))$ is the adjugate of $M(p)$ and $\det(M(p))$ is the determinant of $M(p)$. Therefore, the immersed model matrices are given as

$$\begin{aligned} A_{im}(p) &= d_{\mathfrak{I}\mathfrak{I}} \left(A_{\mathfrak{M}\mathfrak{M}} - A_{\mathfrak{M}\mathfrak{I}} \frac{\mathrm{adj}(A_{\mathfrak{I}\mathfrak{I}})}{\det(A_{\mathfrak{I}\mathfrak{I}})} A_{\mathfrak{I}\mathfrak{M}} \right), \\ B_{im}(p) &= d_{\mathfrak{I}\mathfrak{I}} \left(B_{\mathfrak{M}} - A_{\mathfrak{M}\mathfrak{I}} \frac{\mathrm{adj}(A_{\mathfrak{I}\mathfrak{I}})}{\det(A_{\mathfrak{I}\mathfrak{I}})} B_{\mathfrak{I}} \right), \\ V_{im}(t) &= d_{\mathfrak{I}\mathfrak{I}} \left(V_{\mathfrak{M}}(t) - A_{\mathfrak{M}\mathfrak{I}} \frac{\mathrm{adj}(A_{\mathfrak{I}\mathfrak{I}})}{\det(A_{\mathfrak{I}\mathfrak{I}})} V_{\mathfrak{I}}(t) \right), \end{aligned} \tag{60}$$

where $d_{\mathfrak{I}\mathfrak{I}}(p)$ is a scalar polynomial to ensure that the matrices for the immersed network model become polynomial

after multiplying the rational matrix $A_{\mathfrak{J}\mathfrak{J}}^{-1}$. Typically, $d_{\mathfrak{J}\mathfrak{J}}(p)$ is set as $\det(A_{\mathfrak{J}\mathfrak{J}})$ if $\det(A_{\mathfrak{J}\mathfrak{J}})$ and $\mathrm{adj}(A_{\mathfrak{J}\mathfrak{J}})$ have no common factors. In cases where a common factor is present, dividing by the greatest common divisor of these elements is required to achieve a uniquely represented immersed diffusively coupled network.

According to [13], we know that to identify the target subnetwork, it is sufficient to measure at least the node signals of the target subnetwork and all neighbor node signals. The immersed network can also be divided into two parts, the target subnetwork part $J = \{j | w_j \in w_J\}$ and the neighbors $D = \{j | w_j \in w_D\}$ with $J \cap D = \mathfrak{M}$. The partition results of the immersed diffusively coupled network are:

$$\underbrace{\begin{bmatrix} \bar{A}_{JJ} & \bar{A}_{JD} \\ \bar{A}_{DJ} & \bar{A}_{DD} \end{bmatrix}}_{A_{im}(p)} \underbrace{\begin{bmatrix} w_J(t) \\ w_D(t) \end{bmatrix}}_{w_{\mathfrak{M}(t)}} = \underbrace{\begin{bmatrix} \bar{B}_J \\ \bar{B}_D \end{bmatrix}}_{B_{im}(p)} r(t) + \underbrace{\begin{bmatrix} \bar{V}_J(t) \\ \bar{V}_D(t) \end{bmatrix}}_{V_{im}(t)}. \quad (61)$$

The partition of the original diffusively coupled network is:

$$\underbrace{\begin{bmatrix} A_{JJ} & A_{JD} & 0 \\ A_{DJ} & A_{DD} & A_{D\mathfrak{J}} \\ 0 & A_{\mathfrak{J}D} & A_{\mathfrak{J}\mathfrak{J}} \end{bmatrix}}_{A(p)} \underbrace{\begin{bmatrix} w_J(t) \\ w_D(t) \\ w_{\mathfrak{J}}(t) \end{bmatrix}}_{w(t)} = \underbrace{\begin{bmatrix} B_J \\ B_D \\ B_{\mathfrak{J}} \end{bmatrix}}_{B(p)} r(t) + \underbrace{\begin{bmatrix} V_J(t) \\ V_D(t) \\ V_{\mathfrak{J}}(t) \end{bmatrix}}_{V(t)}, \quad (62)$$

where

$$A_{\mathfrak{M}\mathfrak{M}} = \begin{bmatrix} A_{JJ} & A_{JD} \\ A_{DJ} & A_{DD} \end{bmatrix}, A_{\mathfrak{M}\mathfrak{J}} = \begin{bmatrix} 0 \\ A_{D\mathfrak{J}} \end{bmatrix}, A_{\mathfrak{J}\mathfrak{M}} = \begin{bmatrix} 0 & A_{D\mathfrak{J}} \end{bmatrix},$$
$$B_{\mathfrak{M}} = \begin{bmatrix} B_J \\ B_D \end{bmatrix}, V_{\mathfrak{M}} = \begin{bmatrix} V_J \\ V_D \end{bmatrix}. \quad (63)$$

Remark 14. Notice that $\bar{A}_{JD}(p) = \bar{A}_{DJ}(p)^T$, and we have a relationship as

$$\begin{aligned} \bar{A}_{JJ} &= d_{\mathfrak{J}\mathfrak{J}} A_{JJ}, \quad \bar{A}_{JD} = d_{\mathfrak{J}\mathfrak{J}} A_{JD}, \\ \bar{A}_{DJ} &= d_{\mathfrak{J}\mathfrak{J}} A_{DJ}, \quad \bar{B}_J = d_{\mathfrak{J}\mathfrak{J}} B_J, \end{aligned} \quad (64)$$

where $A_{JJ}(p)$ contains the target subnetwork dynamics of the original model $A(p)$, $A_{JD}(p) = A_{DJ}(p)^T$ contains the dynamics that in the interconnection between the target subnetwork and the neighbor nodes, $B_J(p)$ contains the excitation signal dynamics that entering the target subnetwork. The proof of invariant local dynamics is shown in [13]. By giving a constraint to find the scalar polynomial $d_{\mathfrak{J}\mathfrak{J}}$, the dynamics of the target subnetwork can be identified. Additionally, since users assign node serial numbers to identify a target subnetwork that does not start with the first serial number, it is simple to rearrange the rows and columns to match the target subnetwork nodes and the neighbor measured nodes.

B. The subnetwork identification procedure

Identifying diffusively coupled subnetwork in the frequency domain involves three steps. (1) The first step follows Section III the full network frequency domain identification algorithm, to estimate the FRF and noise covariance of the immersed network. (2) The second step involves using them to estimate the parametric immersed network model, incorporating user-defined constraints. Since the scalar $d_{\mathfrak{J}\mathfrak{J}}$ is unknown, at least one constraint on a non-zero parameter is needed to guarantee a unique solution, which leads to a scaled estimation of the immersed network in comparison to the actual immersed network if that parameter constraint is not equal to the actual one. (3) The last step is to recover the scaled parameters of the set of target subnetwork by giving one known parameter of the original model, such as the parameter in the excitation matrix $B_{\mathfrak{M}}(p)$. For the first two steps, the same procedure can be used directly to identify the scaled immersed network, and the recovery algorithm (immersed back) is introduced in this chapter, in line with [13].

For employing the full network identification algorithm of the first two steps, some conditions similar to condition 1 need to be satisfied. Consider a parametric model set $\mathcal{M}_{im}$ of the immersed network

$$\mathcal{M}_{im} = \{A_{im}(p,\eta), B_{im}(p,\eta), C_{im}(p,\eta), F_{im}(p,\eta)\}, \quad (65)$$

where η collects all the unknown parameters of the model matrices $A_{im}(p), B_{im}(p), F_{im}(p)$ and the transient model $C_{im}(p)$. The data generated immersed network is denoted by $\mathcal{S}_{im} = \{A_{im}^0(p), B_{im}^0(p), F_{im}^0(p)\}$. Furthermore, for the plant model $G_{0_{im}} = \{A_{im}^0(p), B_{im}^0(p)\}$ and $\mathcal{G}_{im} = \{A_{im}(p,\eta), B_{im}(p,\eta), C_{im}(p,\eta)\}$ are the subset of $\mathcal{S}_{im}$ and $\mathcal{M}_{im}$, respectively.

Condition 2. The estimated parameters $\hat{\eta}$ give a consistent estimate of the immersed network if the following conditions hold [13].

1) The true system is in the model set: $\mathcal{S}_{im} \in \mathcal{M}_{im}$
2) The polynomials $A_{im}(p,\eta)$ and $B_{im}(p,\eta)$ are left-coprime.
3) There exists a permutation matrix P that leads to $[A_{im,0}\ A_{im,1}\ \cdots\ A_{im,n_{aim}}\ B_{im,0}\ \cdots\ B_{im,n_{bim}}]P = [D(\eta)\ U(\eta)]$ with $D(\eta)$ square, diagonal and full rank.
4) There exists at least one excitation signal, $K \geq 1$.
5) There exists at least one constraint on the parameters of $A_{im}(p,\eta)$ and $B(p,\eta)$ that ensures that $\Gamma\eta = \gamma \neq 0$.
6) The system is excited at all frequencies of interest.

Remark 15. The proof is given in [13]. However, since the noise model is not parameterized in this frequency domain identification algorithm, the noise model is independent of the plant model. We only care about the parameters of the network plant, so the first condition for the consistent estimate can be relaxed to $G_{0_{im}} \in \mathcal{G}_{im}$.

The target subnetwork can be estimated from the identified immersed network. Given the estimated result of the immersed

network $\hat{A}_{im}(\hat{\eta})$ and $\hat{B}_{im}(\hat{\eta})$, from the relation in (61) and (64) the estimated target subnetwork as

$$\begin{aligned}\hat{A}_{JJ}(\hat{\eta}) &= \alpha d_{\mathfrak{J}\mathfrak{J}} A_{JJ},\\ \hat{A}_{JD}(\hat{\eta}) &= \alpha d_{\mathfrak{J}\mathfrak{J}} A_{JD},\\ \hat{B}_{J}(\hat{\eta}) &= \alpha d_{\mathfrak{J}\mathfrak{J}} B_{J},\end{aligned} \tag{66}$$

where α is the unknown scaling factor that $\hat{A}_{JJ}(\hat{\eta}) = \alpha \bar{A}_{JJ}$, $\hat{A}_{JD}(\hat{\eta}) = \alpha \bar{A}_{JD}$ and $\hat{B}_{J}(\hat{\eta}) = \alpha d_{\mathfrak{J}\mathfrak{J}} B_J$. Substitute the first relation $A_{JJ}^{-1}\hat{A}_{JJ} = \alpha d_{\mathfrak{J}\mathfrak{J}}$ into the rest two equation of (66),

$$\begin{aligned}A_{JJ}\hat{A}_{JD}(\hat{\eta}) - \hat{A}_{JJ}(\hat{\eta})A_{JD} = M_3 = 0,\\ A_{JJ}\hat{B}_{J}(\hat{\eta}) - \hat{A}_{JJ}(\hat{\eta})B_{J} = M_4 = 0.\end{aligned} \tag{67}$$

Collected these two equations into the matrix form,

$$Q_3\theta_3 = 0, \tag{68}$$

where Q_3 collects all the frequency data for the known estimated part $\hat{A}_{JJ}(\Omega_k, \hat{\eta})$,$\hat{A}_{JD}(\Omega_k, \hat{\eta})$ and $\hat{B}_{J}(\Omega_k, \hat{\eta})$ and θ_3 collects all the parameters of the target subnetwork A_{JJ}, A_{JD} and B_J. Here, since $\hat{A}_{im}(\hat{\eta})$ and $\hat{B}_{im}(\hat{\eta})$ is already estimated as a known part, the estimation problem leads to a convex optimization problem. Employing the same idea as before, use the Lagrangian and KKT conditions to solve the least squares problem with constraint,

$$\begin{aligned}\hat{\theta}_3 = \min_{\theta_3} \theta_3^H Q^H(\hat{\eta})Q(\hat{\eta})\theta_3,\\ \text{subject to } \Gamma_3\theta_3 = \upsilon_3.\end{aligned} \tag{69}$$

The estimated results are given:

$$\begin{bmatrix}\hat{\theta}_3\\ \hat{\lambda}_3\end{bmatrix} = \begin{bmatrix}2Q_3(\hat{\eta})^H Q_3(\hat{\eta}) & {\Gamma_3}^T\\ \Gamma_3 & \mathrm{O}\end{bmatrix}^{-1}\begin{bmatrix}\mathrm{O}\\ \upsilon_3\end{bmatrix}, \tag{70}$$

Where O are the zero matrices with the corresponding suitable dimension related to the constraint setting. The structure of the regression matrix Q_3 and the parameter vector θ_3 are given in the Appendix VII-E.

Setting the constraint matrix for identifying the subnetwork for the diffusively coupled network model is similar to the previous full network case. We set the constraint to keep the symmetric structure of A_{JJ} to follow the relation,

$$a_{ij}(p) - a_{ji}(p) = 0 \text{ with } i, j \in J \tag{71}$$

and also at least one parameter of $a_{J_{\bullet_{ij}}}(p)$ or $b_{J_{ij}}(p)$ needs to be known as the constraint to obtain the consistent estimate. The constraint selection matrix Γ_3 and the constraint vector υ_3 are given similarly as Γ and υ presented in Appendix VII-C, but the parameters are collected in a different sequence θ_3 as shown in Appendix VII-E. Thus, the approach involves applying a similar constraint method, but adapting the constraint settings to match the parameter sequence in θ_3.

C. *Subnetwork Identification Procedure Flow*

To conclude the subnetwork identification procedure, the following flow is given:

1) **Non-parametric Identification for immersed network**
   - **Input:** The input and output signals data set, frequency range.
   - **Process:** Using LPM method to estimate $\hat{\Theta}$ in (19) and noise covariance $\hat{C}_V$ in (22).
   - **Output:** The estimated FRF $\hat{G}_{im}$ from $\hat{\Theta}$ and the noise covariance $\hat{C}_{V_{im}}$.
2) **Parametric Identification for immersed network**
   - **Input:** The user-defined parametric model order $n_{a_{im}}$ of $A_{im}(p,\eta)$, $n_{b_{im}}$ of $B_{im}(p,\eta)$, $n_{c_{im}}$ of $C_{im}(p,\eta)$; at least one constraint on non zero parameter e.g., $b_{ij,l} = 1$; input/output frequency data; $\hat{G}_{im}$ and $\hat{C}_{V_{im}}$.
   - **Process:** If $n_{c_{im}} > -1$, using the input/output data criterion shown in Section III-B2 optimize the weighted least squares following (37) with constraint (44) to estimate the parameters $\hat{\eta}$, otherwise using the FRF data criterion shown in Section III-B3 optimize the weighted least squares following (41) with constraint (44) to estimate parameters $\hat{\eta}$.
   - **Output:** Estimation of the immersed network $\hat{A}_{im}(\hat{\eta})$ and $\hat{B}_{im}(\hat{\eta})$.
3) **Target Subnetwork Identification**
   - **Input:** $\hat{A}_{im}(\hat{\eta})$, $\hat{B}_{im}(\hat{\eta})$, topology of $A_{J_\bullet}$, at least one known non-zero coefficient of $A_(p,\theta_3)$ or $B_(p,\theta_3)$, the order of target subnetwork n_a of $A_(p,\theta_3)$, n_b of $B(p,\theta_3)$.
   - **Process:** Obtain the target subnetwork parameters $\hat{\theta}_3$ by optimizing the least squares with the constraint in (70).
   - **Output:** Estimation of the target subnetwork parameters $A_{JJ}(\hat{\vartheta}_3)$, $A_{JD}(\hat{\theta}_3)$ and $B_J(\hat{\theta}_3)$.

Remark 16. Note that the topology information used in the last step input is from the first step topology output of the original network, as immersion will not change the topology of the target subnetwork. Although topology information is not essential for identification, it provides information on the locations of zero terms in $A(p)$ and $A_{im}(p)$. Incorporating this into the constraints can enhance convergence speed and reduce computational complexity. Moreover, this improves the precision of the estimation by reducing the degrees of freedom in the parameter space, which can reduce bias and variance. In addition, this identification flow is fully in the frequency domain to identify the frequency-domain model. Therefore, this flow works the same for the identification of the continuous-time model and the discrete-time model.

D. *Simulation results*

To test the frequency domain immersed network identification algorithm, we used the same seven-node

RLC network as shown in Fig. 5 with the same components as shown in Table I and Table II to simulate. The orders of the parametric subnetwork model are $n_a = 2$ and $n_b = 1$. The CT and DT models share the same order settings. The constraints for the identification of the parametric immersion network of the procedure flow are given as the known coefficient on the input node, e.g., $b_{11,1} = 1$ for exciting the first node, and the known unconnected parts of the network represented as zero terms $A_{zeros_{im}}$ in the matrix $A_{im}(p)$ for the CT model ($B(q^{-1})$, $A(q^{-1})$ for the DT model). The constraints for the identification of the target subnetwork of the procedure flow are given as the known input matrix $B(p)$, the known unconnected parts of the network represented as zero terms A_{zeros} in the matrix $A(p)$ for the CT model ($B(q^{-1})$, $A(q^{-1})$ for the DT model).

Example 3. Suppose the target subnetwork is the interconnection between node 1 and node 2. We measured nodes 1, 2, 3, 4, and 5, while nodes 6 and 7 are immersed. Give the excitation signal $r(t)$ as independent zero-mean white noise with variance $\sigma_r^2 = 1$ entering only node 1, and the normal distributed zero-mean white noise as noise signal $e(t)$ with variance $\sigma_e^2 = 1$ entering all 7 nodes. The sampling frequency is set at 20000 Hz to cover all the dynamics of the components, and the identification frequency band is set between $f_{min} = 500$ Hz and $f_{max} = 6000$ Hz to promote estimation performance. Furthermore, the orders of the immersed parametric network model are defined as $n_{a_{im}} = 3$ of $A_{im}(p, \eta)$, $n_{b_{im}} = 2$ of $B_{im}(p, \eta)$, and $n_{c_{im}} = 0$ for the transient model (CT and DT share the same order settings). Here, the criterion based on the input/output data shown in Section III-B2 is used. Similarly to Example 1, we generated a set of experiments with lengths of data N. The choice of N and the corresponding set are shown in Table III. Each set of experiments includes 50 Monte Carlo runs with independent excitation and noise signals. The results are shown in Fig. 8 for the continuous-time model and in Fig. 9 for the discrete-time model.

It can also be seen from these figures that, in general, RMSE decreases as N increases, supporting a consistent identification. When the data length N is extended to infinity, the estimated parameters converge to the actual parameters. Note that the central line of the box, which indicates the median RMSE, increases from the second experiment to the third experiment. It seems like convergence around 10^{-5} of the first four experiments. This could be due to the insufficient Monto Carlo runs in each experiment and the data length of 1-4 experiments being insufficient to capture all the model dynamics. Moreover, there is a noticeable decrease in the variability of RMSE (indicated by the height of the boxes) from the first to the last sets, which implies that the estimates become more accurate as the data length increases.

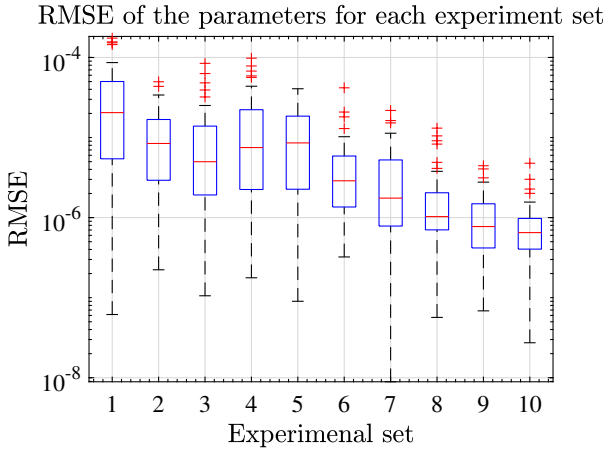


Figure 8: Boxplot of the RMSE of the parameters of the components for each experimental set (CT model immersed 6 and 7)

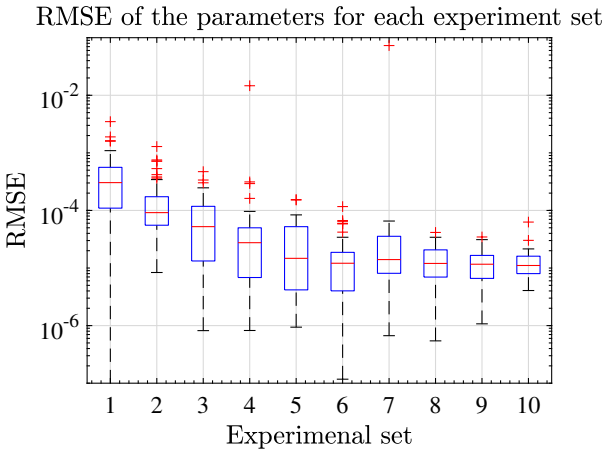


Figure 9: Boxplot of the RMSE of the parameters of the components for each experimental set (DT model immersed 6 and 7)

Example 4. Suppose that the target network in Fig. 5 is the interconnection between node 1 and node 2; we measured the target subnetwork nodes 1 and 2 and their neighbor nodes 3 and 5. Give the excitation signal $r(t)$ as the independent zero-mean white noise with variance $\sigma_r^2 = 1$ entering only node 1, and the normal distributed zero-mean white noise as the noise signal $e(t)$ with variance $\sigma_e^2 = 0.01$ entering all 7 nodes. The sampling frequency is set at 20000 Hz to cover all the dynamics of the components, and the identification frequency band is set between $f_{min} = 500$ Hz and $f_{max} = 6000$ Hz to promote estimation performance. Furthermore, the orders of the parametric immersed network are defined as $n_{a_{im}} = 4$ of $A_{im}(p, \eta)$, $n_{b_{im}} = 3$ of $B_{im}(p, \eta)$, and $n_{c_{im}} = 1$ for the transient model. Here, the criterion based on the input/output data shown in Section III-B2 is used. Similarly to Example 1, we generated a different set

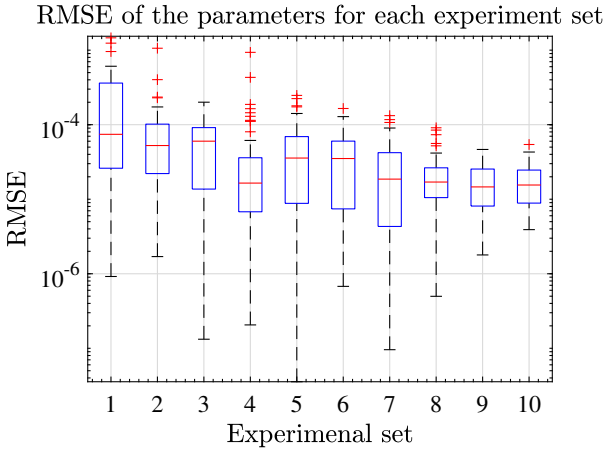


Figure 10: Boxplot of the RMSE of the parameters of the components for each experimental set (CT model immersed 4, 6 and 7)

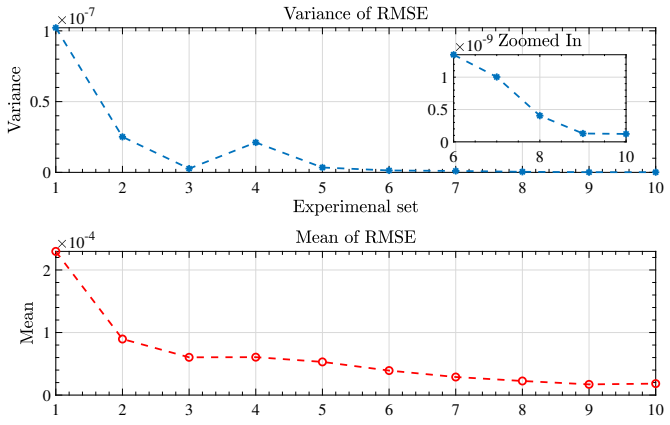


Figure 11: The mean and variance of the RMSE for component parameters across each experimental set (CT model immersed 4, 6, and 7)

of experiments with a difference in the length of the data N. The choice of N and the corresponding set are shown in Table III. Each set of experiments includes 50 Monte Carlo runs with independent excitation and noise signals. The results are shown in Fig. 10 for the continuous-time model with the mean and variance of the RMSE shown in Fig. 11.

From Fig. 10, it can be seen that there is no significant trend to decrease the RMSE of the component value with increasing data length in this case. However, we notice that the number of outliers and the interquartile range decrease with increasing data length. Also, the median value of each box is decreasing, except for the fourth experiment. To check whether the RMSE slowly decreases in this case, the variance and the mean value of the RMSE are provided with each experiment in Fig.11. Fig.11 shows that the mean value of RMSE for each experiment decreases with increasing data length. This trend suggests that the estimation process is consistent and improves accuracy as more data are collected. The variance is mainly decreasing, except for the fourth experiment. This might be due to numerical problems or overfitting of the data. The variance value is already very small in this case and will not continue to increase in the coming experiments. Generally, the trends are good in both the variance and mean plots, indicating a consistent estimation in this example. Notice that the experiment result for the discrete-time model is not given due to the unstable calculation of the discrete-time model in Matlab. Especially to calculate the discrete-time inverse matrix $A_{im}(p)$ during the simulation, the more components are immersed, the higher the value and the orders will be in $A_{im}(p)$. Furthermore, the higher the sampling frequency we choose to discretize the model, the value in $A_{im}(p)$ will also increase. The discrete-time matrix $A_{im}(p)^{-1}$ calculated by Matlab is unstable in this case. Thus, the result is not shown. In conclusion, the second task of this project is achieved, a consistent frequency domain identification method is developed for the diffusively coupled subnetwork, and the consistency of it is shown by the experiments.

V. The In-Circuit Testing Procedure with Using Frequency Domain Identification Algorithm

The preceding sections outline the frequency domain identification algorithm applied to the full network and subnetwork within a diffusively coupled framework. Some experiments also showed that this algorithm can be used to identify the components of the RLC circuit. The last task of this project is to develop a testing procedure for applying the frequency domain diffusively coupled network identification algorithm to ICT simulation and testing it with different defects. This section presents an in-circuit testing procedure with the frequency domain identification algorithm consisting of three parts: the data level, the feature level, and the parameter level. The flow diagram of this procedure is shown in Fig. 12. As is well known, an ideal short circuit is a connection between two nodes that causes them to have the same voltage signal. Some time-domain data characteristics, such as magnitude and power, can be extracted from the data set to judge similar voltage signals if there is probably a short circuit. For the feature level, the frequency domain features: FRF and resonance peaks will be discussed to detect faults. The parameter level focuses on comparing the estimated parameters of the components and the corresponding parameters in a healthy model. Finally, 4 experiments will be shown to demonstrate the ability of this ICT procedure.

A. *Data level*

Suppose we have a data set that includes the input signals $r(t)$ and node signals $w(t)$. The first thing we could check is the time-domain data set. Observing the data of the node signal can provide valuable information. For example, as shown in Fig. 12, a node signal (voltage) that is approximately

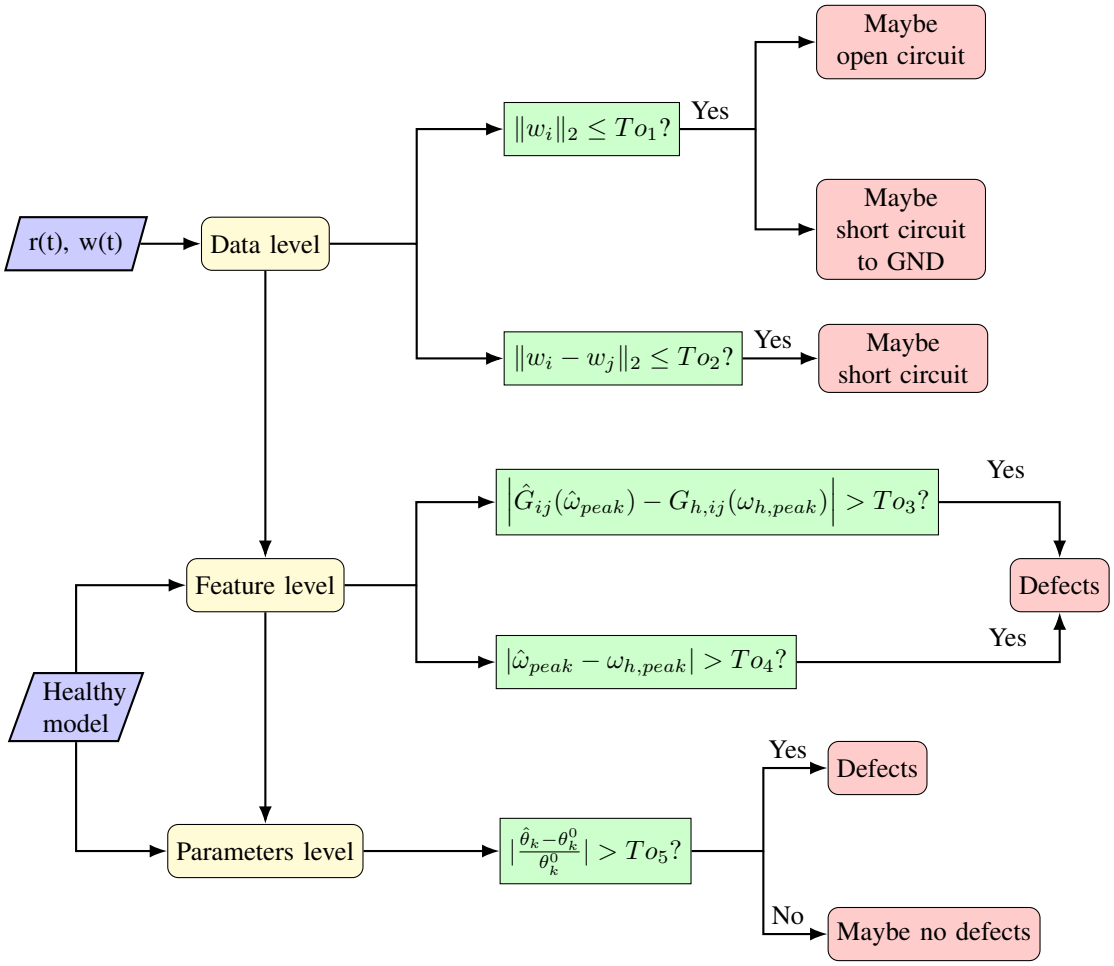


Figure 12: ICT with using frequency domain identification flow diagram ($To_1, To_2, To_3, To_4, To_5$ are the tolerance values, $\hat{G}_{ij}$ is the FRF estimationm, $G_{h,ij}$ is the healthy model FRF, $\hat{\omega}_{peak}$ is the resonance peak frequency in FRF estimation, $\omega_{h,peak}$ is the resonance peak frequency in healthy model FRF)

zero $w_i \approx 0$ could indicate that the node is grounded (Short Circuit to GND) or disconnected from the electronic network (open circuit); two signals that are approximately the same $w_i \approx w_j$ could indicate that a short circuit is connected between these two nodes.

Various statistical and numerical methods can be used to check whether $w_i \approx 0$. For example, calculating the mean absolute value (MAV) and the root mean square (RMS) of the data set for the $i-th$ signal w_i to see how small they are, or setting a threshold value ϵ_{th} if all elements of the set are within $0 \pm \epsilon_{th}$ then w_i can be considered to be nearly equal to 0. In this project, we calculate the 2-norm of the data set for $i-th$ signal w_i that gives the magnitude of the data vector. Then, we take $\log_{10}$ of each 2-norm for the corresponding node signal data set and compare it with the tolerance value. For the signals $w_i \approx w_j$, we also take $\log_{10}$ of each 2-norm for the corresponding node signals and calculate the difference of each other if $|\log_{10}\|w_i\|_2 - \log_{10}\|w_j\|_2| < 0.001$ then we consider these two signals to be approximately the same.

Notice that there is no conclusion on this level about whether there is a fault. The data level gives some information that can be used to help users find faults and isolate faults in the following steps.

B. *Feature level*

Following the frequency domain identification algorithm, the Fourier transform is applied to transform the signals to the frequency domain after collecting the time domain data. The features of the signal can be obtained, for example, the amplitudes, phases, and spectrum. When observed characteristics are compared with healthy behaviors, differences are considered analytical symptoms to detect the fault. Here, instead of comparing the signal features directly, we consider one step further. Since the first step of the frequency domain identification algorithm shown in Section III is to estimate the FRF and the noise covariance, we can use the FRF estimation as the model features of the RLC circuits and compare it with the FRF of the healthy

model for the healthy RLC circuits. Changes in the magnitude of the FRF can indicate potential faults or changes in the dynamics of the system. For example, an RC series circuit with a grounded capacitor behaves as a low-pass filter; if the faults appear in the circuit, the cut-off frequency and the resonances of the FRF would be changed.

To check whether the FRF is healthy or not, we can directly calculate the RMS of $\hat{G}_{ij} - G_{h,ij}$ ($\hat{G}_{ij}$ is the estimation of the FRF between the input node i and the output node j, $G_{h,ij}$ is the healthy FRF between the input node i and the output node j). In this project, we extract the resonance peaks, calculate the cut-off frequency of each FRF, and compare the estimated FRF with the healthy FRF 95% confidence interval.

An example is shown in Fig. 13. This figure shows the FRF estimation of the CT seven-node defect model with the true FRF of the healthy model. The defect model is given as three faults appear in the circuit simultaneously, including a dynamic change in R_{13} from 100 Ω to 500 Ω, an open circuit in L_{34} and a short circuit in R_{57}. The open circuit here is considered the inductor with an infinitely large inductance $L_{34} = \infty$ mH, $\frac{1}{L_{34}} = 0$, and the short circuit here is considered a resistor with a very small resistance as $R_{57} = 0.001$ Ω. Using the same experimental setting as in Example 2, give the excitation signal $r(t)$ as the independent zero mean white noise with variance $\sigma_r^2 = 1$ entering only node 6, and the normal distributed zero-mean white noise as the noise signal $e(t)$ with variance $\sigma_e^2 = 1$ entering all nodes. The frequency band is set between 500 Hz and 6000 Hz.

As can be seen in Fig. 13, the FRFs are all outside the $\pm 5\%$ bound of the true FRFs, except for the FRF from nodes 6 to 6. The resonance peaks estimation also does not correspond to the peaks of true FRFs, except for the FRF from node 6 to node 6 and from node 6 to node 1. This happens because the faulty dynamic might be filtered out of the FRF estimation, and the estimation is not sensitive enough to capture it. The feature level information can detect faults by estimating the FRF and comparing the features with the true FRF.

Moreover, from many experiments, we can observe that the defects show a diffusion phenomenon in the network, like throwing a stone into a calm lake. For input/output nodes located away from the defect, the variation in the FRF between faulty and healthy models is less than that for nodes closer to the defect. For example, the 7-node network model in Fig. 5 with only an open circuit in L_{34}. The full excited and full measurement setup with the same types of excitation signals and noise signals as in Example 2 is given. The frequency band is set between 500 Hz and 6000 Hz. The FRF estimations of this CT seven-node faulty model with the true FRFs of the healthy model are shown in Fig. 23 in Appendix VII-F. The estimations of FRFs from nodes 2 to 4, nodes 3 to 4, and nodes 4 to 4 (and their symmetric FRFs)

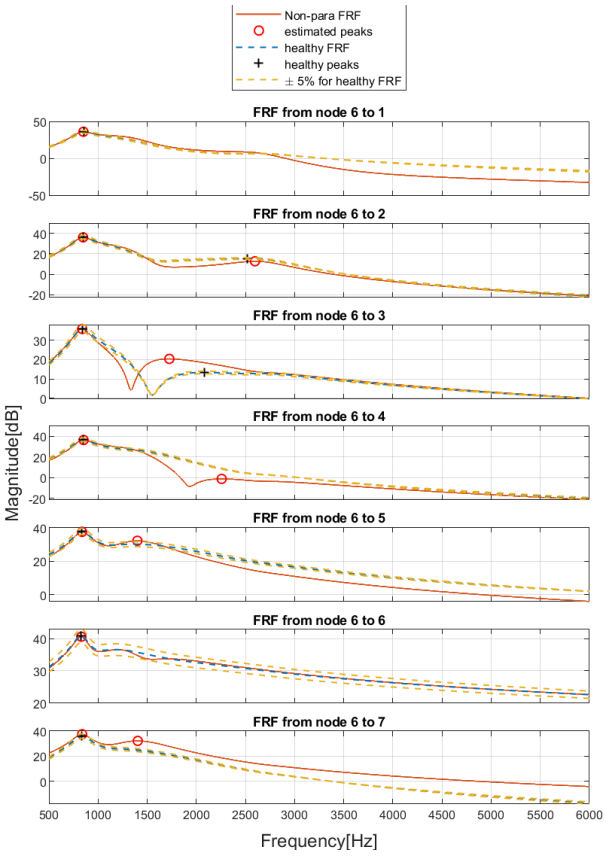


Figure 13: FRF estimation of defect model with the true FRF of the healthy model (CT model)

show the greatest difference from the healthy FRFs, while the estimations of FRFs such as from nodes 1 to 1, 1 to 2, 5 to 5, 5 to 6, 6 to 6, 7 to 7 (and their symmetric FRFs) are almost the same as the healthy model and stay in the $\pm 5\%$ bound of the healthy FRFs. According to this FRF information, the user can speculate that a defect might occur in the component that connects to node 4. This phenomenon can be explained as the faulty dynamic might be progressively filtered as the signals pass through an increasing number of components. Therefore, by utilizing FRF information, users can narrow down the potential fault area of the network when there are few faults.

C. *Parameters level*

At the parameters level, the idea is to identify all the component values and compare them with the true value so that fault isolation and identification can be done simultaneously. The relative parameter error of each component is calculated, which is defined as

$$\text{RPE} = \frac{\hat{\theta}_{comp_i} - \theta^0_{comp_i}}{\theta^0_{comp_i}}, \quad \text{for } i\text{-th component}, \tag{72}$$

where $\hat{\theta}_{comp_i}$ is the parametric estimation of i-th component and $\theta^0_{comp_i}$ is the real parameter value of the i-th component. This relative parameter error indicates the percentage of parametric changes. From the literature, we know that the measurement value of the components should also remain in $\pm 5\%$ change due to measurement and manufacturing errors (e.g., temperature will also influence the electronic components). Therefore, the tolerance value here for defining the estimation of the relative error of healthy components is given as $Tol_5 = 0.05$.

D. *Simulation results*

This section presents the results of 4 simulated experiments conducted to assess the capability of the ICT procedure using the frequency domain identification algorithm for diffusively coupled networks. It is important to note that this project focuses mainly on three common types of defects in PCBA production: open circuits, short circuits, and changed dynamics, as discussed in Section I-E. Extensive experimentation has revealed that short circuits pose the greatest challenge in analyzing the estimation results. This is because a short circuit behaves like a minuscule resistor between two nodes, introducing the term $\frac{1}{R_{short}}$, which tends to infinity and consequently affects optimization calculations in parametric identification. Moreover, the occurrence of short circuits, especially alongside parallel paths, results in predominantly disturbance signals in these routes, complicating the identification of parallel dynamics due to the weak disturbance signals. In this context, our focus is on scenarios where short circuits occur in single connections without any parallel path and where one of the connected nodes is grounded with the measurement components. We will refer to this situation as the short circuit between the edge nodes of the network.

The experiment simulation procedure is given as follows:

1) Using defect model $\mathcal{M}_{defect}$ to generate the data set, using the band limit measurement setting if the model is in continuous time and the first-order hold measurement setting for discrete time.
2) Using this data set and the information of the healthy model $\mathcal{M}_{healthy}$ (e.g., topology and order of the model) to identify the network model $\hat{\mathcal{M}}_{defect}$.
3) Using the list of components of the healthy model to extract the value of the components $\hat{\theta}_{comp}$ from the estimation of the defect model $\hat{\mathcal{M}}_{defect}$.
4) Calculate the RPE with $\theta^0_{comp_i}$ as the parameter for each healthy component.

Remark 17. All experiments in this section follow this setting: the sampling frequency (f_s) is set at 20000 Hz, and the identification frequency band is set between $f_{min} = 500$ Hz and $f_{max} = 6000$ Hz. The data length N is 20000.

Example 5. This example is set to show the fault detection and diagnosis of the full network with dynamic changes, open circuit, and short circuit. The simulation results of Example 2 show that only one excited node can identify the full dynamics of the network. Using the same experimental setting as in Example 2, given the excitation signal $r(t)$ as the independent zero mean white noise with variance $\sigma_r^2 = 1$ entering only node 6, and the normal distributed zero-mean white noise as the noise signal $e(t)$ with variance $\sigma_e^2 = 1$ entering all nodes. The defect model is given as three faults appear in the circuit simultaneously, including a dynamic change in R_{13} from 100 Ω to 500 Ω, an open circuit in L_{34} and a short circuit in R_{57}. The orders of the parametric model are $n_a = 2$, $n_b = 1$ and $n_c = 1$. The RPE of the estimation results for the CT model and the DT model are shown in Fig. 14 and Fig. 15, respectively. The real parameter of each component and the mean estimated parameter values of 50 Monte Carlo runs (MC) are shown in Table IV, where ***Idx*** is the index number corresponding to the number of components in Fig. 14 and Fig. 15, Hvalue is the healthy (ideal) value of the component in the healthy model, TValue is the true value of the component in the defect model, CT est is the mean estimated value based on the CT model for 50 MC, and DT est is the mean estimated value based on the DT model for 50 MC.

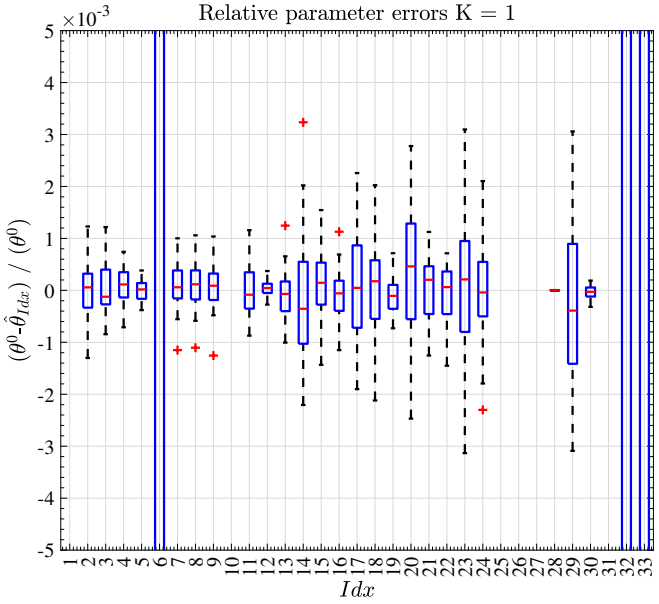


Figure 14: Relative parameters errors for the 7 node defect model (CT model for Example 5)

As can be seen in Fig. 14 and Fig. 15, most of the relative parameter errors stay between $\pm 4 \times 10^{-3}$ and $\pm 5 \times 10^{-2}$, respectively, except for the components R_{13}, R_{57}, L_{34}, C_{50}, R_{50}, L_{50}, C_{70}, R_{70}, L_{70}, respectively. The open circuit in L_{34} and the changed dynamic in R_{13} faults can be directly identified, and the mean estimated value of $\hat{L}_{34}$ increases to a high value or inf ($\frac{1}{\hat{L}_{34}} \approx 0$) and the mean estimated value of $\hat{R}_{13} \approx 500 \neq 100$ Ω. However, for the short circuit in R_{57},

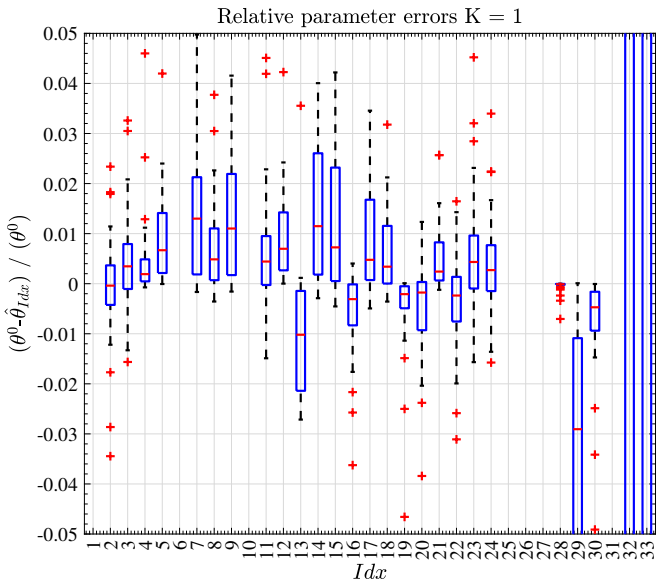


Figure 15: Relative parameters errors for the 7 node defect model (DT model for Example 5)

Table IV: Seven-node defect Network Component Values

Idx	Comp	HValue	TValue	CT est	DT est	Unit
1	R_{13}	100	500	500.0936	509.2552	Ω
2	R_{23}	200	200	200.0136	199.2637	Ω
3	R_{34}	150	150	150.0022	152.8268	Ω
4	R_{36}	180	180	180.0158	181.6703	Ω
5	R_{56}	160	160	160.0019	164.4294	Ω
6	R_{57}	120	0.001	-1.2578e+3	-0.0144	Ω
7	L_{12}	5	5	5	5.2	mH
8	L_{23}	10	10	10	10.1	mH
9	L_{25}	15	15	15	15.4	mH
10	L_{34}	12	Inf	Inf	2252.5	mH
11	L_{45}	20	20	20	20.5	mH
12	L_{56}	13	13	13	13.4	mH
13	C_{10}	2	2	1.9998	1.9302	μF
14	R_{10}	500	500	499.9092	518.7463	Ω
15	L_{10}	18	18	18	18.7	mH
16	C_{20}	2	2	1.9998	1.9840	μF
17	R_{20}	500	500	500.0727	500.6553	Ω
18	L_{20}	18	18	18	18.3	mH
19	C_{30}	2	2	1.9999	1.9811	μF
20	R_{30}	500	500	500.1791	495.8416	Ω
21	L_{30}	18	18	18	18.2	mH
22	C_{40}	2	2	2	1.9661	μF
23	R_{40}	500	500	500.0778	514.5921	Ω
24	L_{40}	18	18	18	18.3	mH
25	C_{50}	2	2	4	4.0298	μF
26	R_{50}	500	500	250.0037	247.4700	Ω
27	L_{50}	18	18	9	8.9	mH
28	C_{60}	2	2	2	1.9992	μF
29	R_{60}	500	500	499.8598	443.8767	Ω
30	L_{60}	18	18	18	17.7	mH
31	C_{70}	2	2	-2.4139e-6	-1.3921e-1	μF
32	R_{70}	500	500	3.5523e+8	-8.0972e+3	Ω
33	L_{70}	18	18	-2.1208e+6	-252.9	mH

the estimated value is not close to the resistance of the short circuit that we set as 0.001, and the grounded measurement components connected to nodes 5 and 7 are not correctly identified. One reason might be due to the small resistance value $R_{short57} = 0.001$ of the short circuit, the corresponding terms $A_{57,1} = A_{75,1} \approx A_{55,1} \approx A_{77,1}$ in the first order of the polynomial matrix $A(p)$ in (1), and these terms are significantly larger than the other terms of the matrix, which could significantly increase the condition number of the matrix and make it difficult to identify. However, the physical dynamics changed by the short circuit in the network and the corresponding solution to identify them still need further research in future work. We have also done similar experiments in the RLC network of 3 nodes and 10 nodes RLC network; the short circuit that appears in the edge nodes of the network will only cause an incorrect estimation of the short circuit itself and the measurement components connected to it. The consistent estimation still holds for the other components. Moreover, if a short circuit occurs, we can detect similar two-node signals from the data level, which also helps the user isolate the short-circuit defect. From these experimental results, biases occur in the relative parameter errors, and there are more outlier values for the DT model compared to the CT model estimation. This small bias of approximately $\pm 1\%$ from the true values in the defect model can be attributed to the insufficiently high sampling frequency and the data length chosen here being not sufficient to use the DT model.

Example 6. The ability to detect and diagnose faults for multiple open circuits and dynamic changes that occur simultaneously in a more complex network will be shown in this example. This example will be shown in a 10-node RLC network as in Fig. 16. The measurement components are the same as in Table I, and the interconnection components of this network are shown as the HValue in Table V. The topology information of this complex 10-node circuit is clearly shown in Fig. 17, where the red path means that there exists one component on the path, two parallel components on the blue path, and three parallel components on the yellow path; the yellow circle is not the self-loop but represents the grounded path. Given the changed dynamics in R_{13}, R_{36}, R_{89}, L_{25}, open circuit in R_{45} and L_{56}. Using the excitation signal $r(t)$ as independent zero-mean white noise with variance $\sigma_r^2 = 1$ entering only node 3, and normal distributed zero-mean white noise as noise signal $e(t)$ with variance $\sigma_e^2 = 1$ entering all nodes. The orders of the parametric model are $n_a = 2$, $n_b = 1$ and $n_c = 1$. The RPE of the estimation results for the CT model and the DT model are shown in Fig. 18 and Fig. 19, respectively. The real parameter of each component and the mean estimated parameter values of 50 MC are shown in Table V, where the information on the estimation measurement components is not shown.

As shown in Fig. 18 and Fig. 19, most of the relative

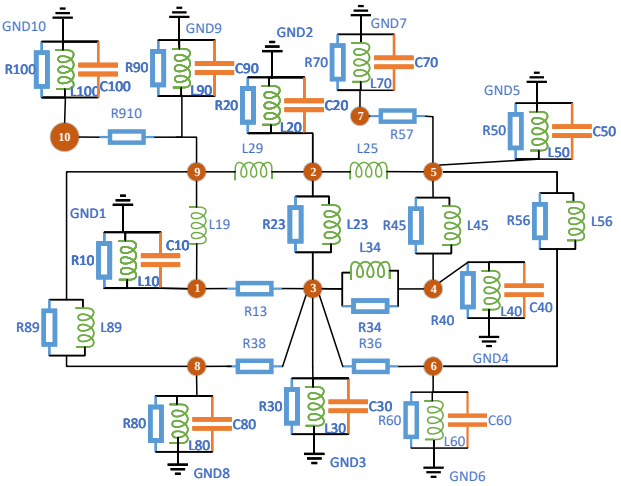


Figure 16: A ten-node RLC network Example with inductors (L_{jk}), resistors (R_{jk}), capacitors (C_{jk}) and ground nodes(GND_j)

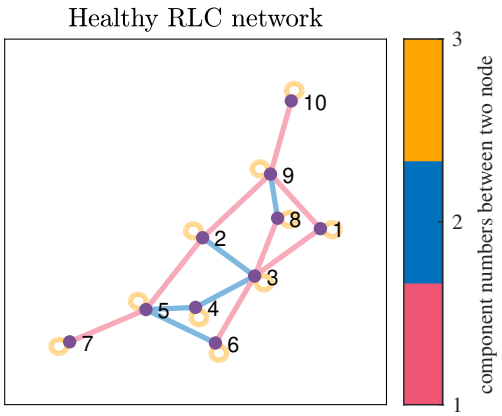


Figure 17: The topology of the 10 nodes RLC circuits (red, blue, and yellow path means that there exists 1, 2, 3 components, respectively)

Table V: Ten-node defect Network Component Values

Idx	Comp	HValue	TValue	CT est	DT est	Unit
1	R_{13}	100	200	200.0376	200.1912	Ω
2	R_{23}	200	200	200.0774	200.1183	Ω
3	R_{34}	150	150	149.9639	150.0634	Ω
4	R_{36}	180	500	499.2765	497.4794	Ω
5	R_{45}	350	Inf	Inf	Inf	Ω
6	R_{38}	180	180	179.9800	180.1738	Ω
7	R_{56}	160	160	159.7945	159.1906	Ω
8	R_{57}	120	120	120.3104	120.5180	Ω
9	R_{89}	160	500	500.1714	500.1982	Ω
10	R_{910}	120	120	119.9661	120.2332	Ω
11	L_{19}	5	5	5	5	mH
12	L_{29}	3	3	3	3	mH
13	L_{23}	10	10	10	10	mH
14	L_{25}	15	1	1	1	mH
15	L_{34}	12	12	12	12	mH
16	L_{45}	20	20	20	20	mH
17	L_{56}	13	Inf	Inf	119552.6	mH
18	L_{89}	13	13	13	13	mH

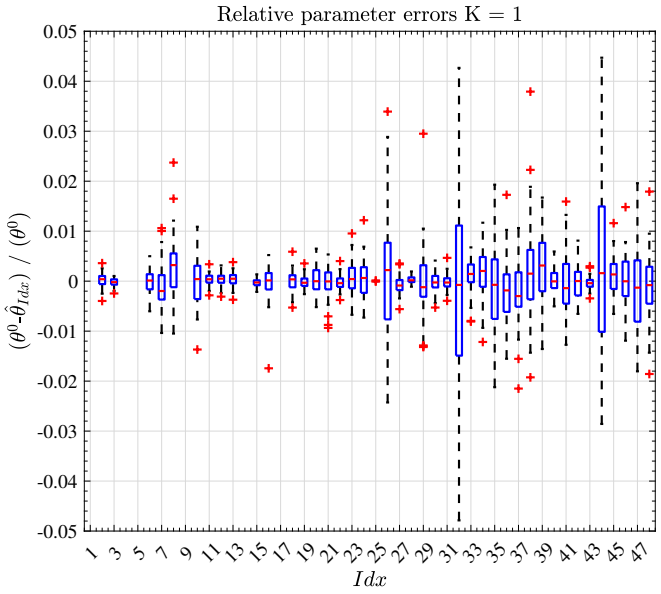


Figure 18: Relative parameters errors for the 10 node defect model (CT model for Example 6)

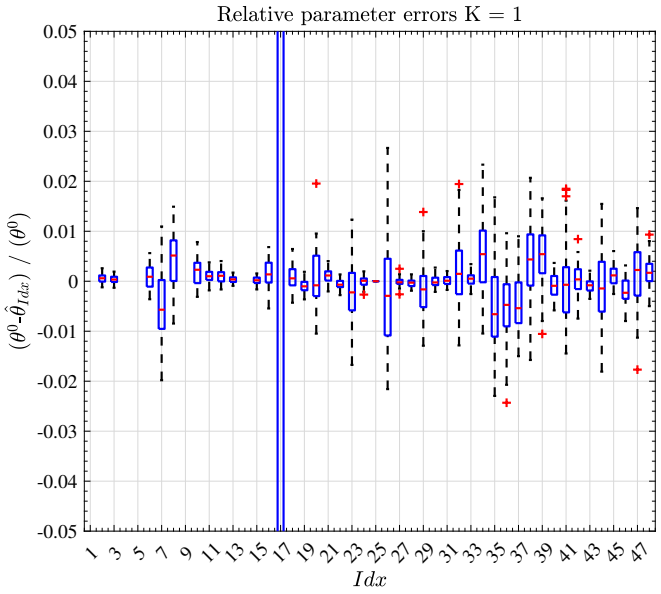


Figure 19: Relative parameters errors for the 10 node defect model (DT model for Example 6)

parameter errors stay between $\pm 5 \times 10^{-2}$ for the CT model experiment and $\pm 3 \times 10^{-2}$ for DT model, except for the components R_{13}, R_{36}, R_{45}, R_{89}, L_{25}, L_{56}. The interquartile range and the whiskers of the box plot of each component value for the CT model experiment are slightly wider than the box plot in the DT model result; more outliers of the CT model experiment are also observed compared to the DT model result. These observations indicate that the variance of the CT result is higher than in discrete time. This is because the length of the data N is not high enough for

this continuous time model, as shown in Example 2 with increasing data length, the interquartile range will be narrower and the number of outliers will also decrease. However, the box plot of the DT experiment has biases in some of the values of the components, e.g. C_{60}, R_{60}, L_{60}, C_{70}, R_{70}, L_{70} ($\boldsymbol{Idx}$.34 to 39). This observation can also be seen in Table V, the mean value of the estimated continuous time 50 MC runs is closer to the true values, and the mean value deviation for the DT experiments is larger compared to the CT value. This is because the sampling frequency of the DT model is not high enough and the discretization causes bias.

Example 7. This example is set up to show the ability to detect and diagnose faults for the subnetwork with short circuit between the edge nodes. This example is based on the same 10-node RLC network as in Fig. 16 but with immersed nodes 4, 6, and 7. The target is to identify defects in the subnetwork consisting of nodes 1, 2, 8, 9, and 10. To identify the dynamics in the subnetwork, we at least need to measure the nodes in the target subnetwork 1, 2, 8, 9, 10, and the neighboring nodes 3 and 5. Given the short circuit in R_{910}, and the dynamic change in L_{19}. Using the excitation signal $r(t)$ as independent zero-mean white noise with variance $\sigma_r^2 = 1$ entering nodes 1 and 9, and the normal distributed zero-mean white noise as noise signal $e(t)$ with variance $\sigma_e^2 = 0.01$ entering all nodes. The orders of the immersed parametric network model are defined as $n_{a_{im}} = 4$ of $A_{im}(p,\eta)$, $n_{b_{im}} = 3$ of $B_{im}(p,\eta)$, and $n_{c_{im}} = 1$ for the transient model. The orders of the parametric model of the target subnetwork are defined as $n_a = 2$, $n_b = 1$. The RPE of the estimation results for the CT model is shown in Fig. 20. The real parameter of each component (with $\boldsymbol{Idx}$ the same index number as in the figures) and the mean estimated parameter values of 50 MC are shown in Table VI.

As can be seen in Fig. 20 and Table VI, most of the relative parameter errors stay between $\pm 2 \times 10^{-2}$, except for the components R_{910}, R_{19}, C_{90}, R_{90}, L_{90}, C_{100}, R_{100}, L_{100}. Similarly to that in Example 5, the short circuit dynamic and measurement components connected to the short circuit nodes cannot be correctly identified. The accurate estimation still holds for the rest of the components. Analysis at the data level reveals a similarity between the signals of nodes 9 and 10, and the parametric results suggest the high possibility of a short circuit occurring between these nodes. Here, the experiment for the DT model is not given, which is because the DT model generated in Matlab with $f_s = 2000$ Hz is unstable (its transfer function exists poles out of the unit circle).

Example 8. This example is designed to show the ability to identify defects in the measurement components of the subnetwork. Based on the same network shown in Fig. 5. Suppose that the target subnetwork is the interconnection between node 1 and node 2, we measured nodes 1, 2, 3, 4,

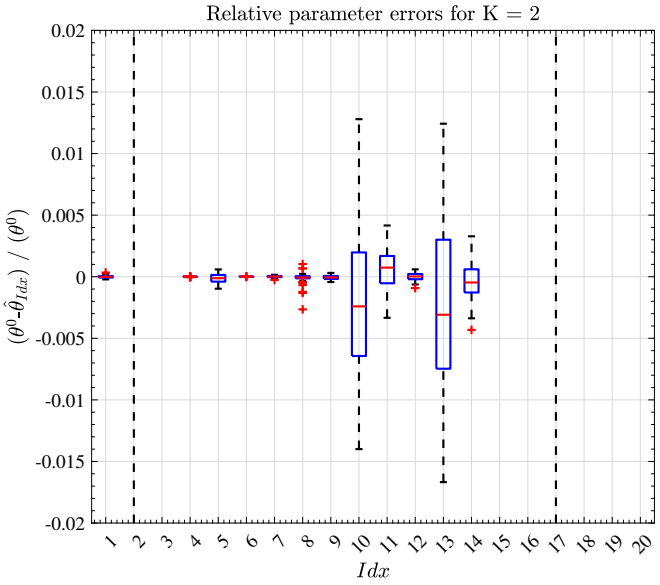


Figure 20: Relative parameters errors for the 10 node immersed defect model (CT model for Example 7)

Table VI: Ten-node immersed defect Network Component Values

$\boldsymbol{Idx}$	Comp	HValue	TValue	CT est	Unit
1	R_{89}	160	160	159.9984	Ω
2	R_{910}	120	0.001	395.0109	Ω
3	L_{19}	5	10	10	mH
4	L_{29}	3	3	3	mH
5	L_{89}	13	13	13	mH
6	C_{10}	2	2	2	μF
7	R_{10}	500	500	500.0081	Ω
8	L_{10}	18	18	18	mH
9	C_{20}	2	2	1.9999	μF
10	R_{20}	500	500	498.7659	Ω
11	L_{20}	18	18	18	mH
12	C_{80}	2	2	2	μF
13	R_{80}	500	500	498.6697	Ω
14	L_{80}	18	18	18	mH
15	C_{90}	2	2	3.4229	μF
16	R_{90}	500	500	221.7812	Ω
17	L_{90}	18	18	-5.9	mH
18	C_{100}	2	2	-6.0975e+14	μF
19	R_{100}	500	500	2.7973e-12	Ω
20	L_{100}	18	18	4.5600e-15	mH

and 5, while nodes 6 and 7 are immersed. Although node 4 could be immersed, as illustrated in Example 4, it is measured in this example to prevent the unstable DT model caused by discretization and to demonstrate the results using the stable DT model. Giving the excitation signal $r(t)$ as independent zero-mean white noise with variance $\sigma_r^2 = 1$ entering only node 1, and the normal distributed zero-mean white noise as noise signal $e(t)$ with variance $\sigma_e^2 = 0.01$ entering all 7 nodes. Defects are given as the dynamic change in the measurement component R_{10} and the interconnected component L_{12}, the open circuit in the measurement component R_{20}. The orders of the immersed parametric network model are defined as

Table VII: Seven-node immersed defect Network Component Values

Idx	Comp	HValue	TValue	CT est	DT est	Unit
1	R_{10}	500	200	200.0108	200.0011	Ω
2	R_{20}	500	Inf	3.9987e+5	4.0970e+5	Ω
3	L_{12}	5	10	10	10	mH
4	C_{10}	2	2	2	2	μF
5	L_{10}	18	18	18	18	mH
6	C_{20}	2	2	2	2.001	μF
7	L_{20}	18	18	18	18	mH

$n_{a_{im}} = 3$ of $A_{im}(p, \eta)$, $n_{b_{im}} = 2$ of $B_{im}(p, \eta)$, and $n_{c_{im}} = 1$ for the transient model. The orders of the parametric model of the target subnetwork are defined as $n_a = 2$, $n_b = 1$. The RPE of the estimation results for the CT model and the DT model are shown in Fig. 21 and Fig. 22, respectively. The real parameter of each component (with $\boldsymbol{Idx}$ the same index number as in the figures) and the mean estimated parameter values of 50 MC are shown in Table VII.

As can be seen in Fig. 21 and Fig. 22, most of the relative parameter errors stay between $\pm 2 \times 10^{-3}$, except for the components R_{10}, R_{20}, and L_{12}, respectively. Table VII shows the values of the components of the healthy model and the components of the faulty model. After 50 MC runs, the mean values of the estimated results for these three components are just in the deviation $5.4 \times 10^{-3}\%$ and $5.5 \times 10^{-4}\%$ with the true faulty values for using the CT model and the DT model, respectively. However, there are still around $1 \times 10^{-3}\%$ deviations for all 50 MC runs with estimated values of the components C_{20} and L_{20}. These

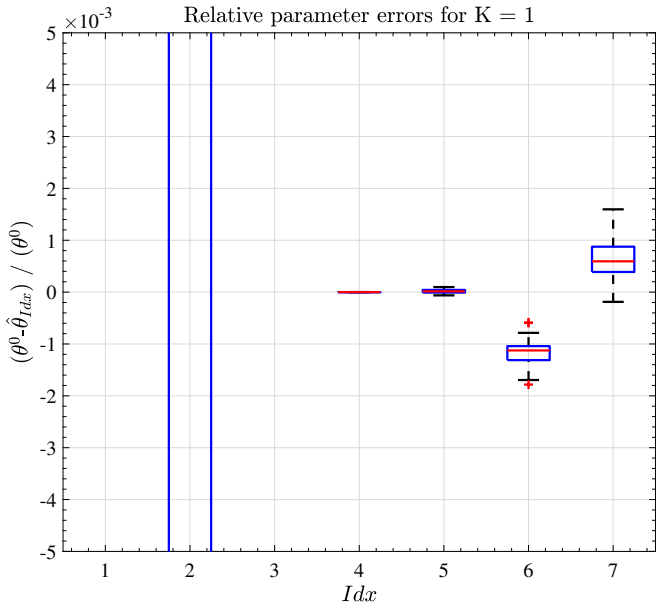


Figure 21: Relative parameters errors for the 7 node immersed defect model (CT model for Example 8)

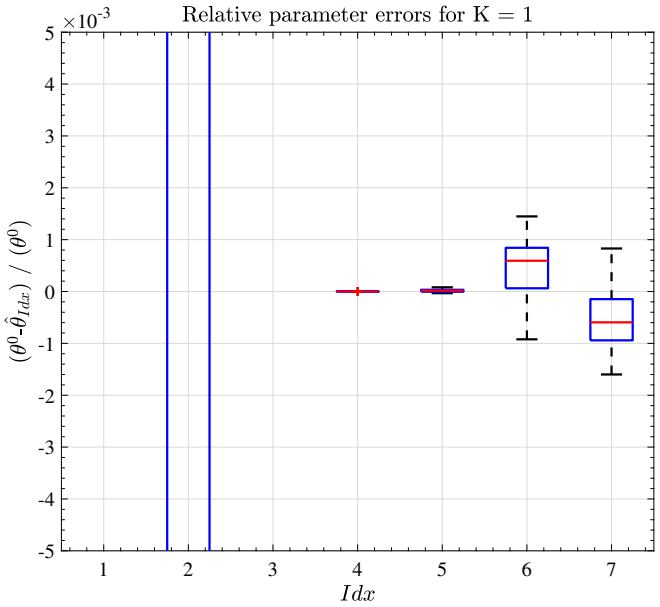


Figure 22: Relative parameters errors for the 7 node immersed defect model (DT model for Example 8)

small biases occur because of the finite data length N and the numerical calculation in Matlab. This experiment shows that even for defects L_{12} that are connected to two faulty measurement components R_{10} and R_{20}, these three faulty components can still be detected and identified simultaneously.

To conclude the experiments we conducted in this section, a list can be made to show the ability of this ICT procedure with applying the frequency domain diffusively coupled network procedure in the ICT simulation.

1) In Example 5, 3 different types of defects occur simultaneously in the 7 node network. Using frequency domain identification for the full network, faults of dynamic changes and open circuits can be detected and diagnosed. However, the short circuit between the edge nodes can only be detected but incorrectly estimated, and the healthy grounded components that connect to the short circuit are also incorrectly estimated. Based on the data-level information, users can infer that the fault is a short circuit.
2) In Example 6, 4 dynamic changes and 2 open circuits occur in the 10 node network simultaneously. Using the frequency domain identification for the full network, although the network is more complex with more defects compared to Example 5, faults of dynamic changes and open circuit can still be detected and diagnosed correctly.
3) In Example 7, 1 short circuit and 1 dynamic change occur in the 5 node subnetwork of the 10 node network. Using frequency domain identification for the subnetwork, dynamic changed fault can be detected and diagnosed correctly. However, the short circuit estimation is

similar to that in Example 5. Short circuit between edge nodes can only be detected but incorrectly estimated, and healthy grounded components that connect to the short circuit are also incorrectly estimated. With the data-level information, users can infer that the fault is a short circuit.

4) In Example 8, 1 open circuit and 1 dynamic change in the interconnection, 1 dynamic change in the grounded component occur in the 2 node subnetwork of the 7 node network. Using frequency domain identification for the subnetwork, dynamic changed in grounded component and open circuit of the subnetwork can be detected and diagnosed correctly, even these three faulty components are connected each other.

Overall, with this three-level ICT procedure that uses the identification of the diffusively coupled network in the frequency domain, three types of common defects can be detected in ICT. Dynamic changes and open circuits can be correctly diagnosed by estimating the parameters. Short circuits that occur between edge nodes can be detected, but cannot be diagnosed at the parameter level. Combining the information from the data level and the feature level, the short circuits can be isolated.

VI. Conclusion and future work

A. Conclusion

This thesis introduces a frequency domain identification method in the dynamic network and uses it for in-circuit testing in PCBA. This approach is based on the diffusively coupled network because it captures the values of the components in the PCBA structurally. A two-step frequency domain identification method is developed to identify components, resulting in consistent estimates that are applicable to the full networks and subnetworks. It has been shown that all the values of the components in a full network can be estimated consistently with a single excitation. Furthermore, the values of the target subnetwork can be consistently estimated even with a single excitation, regardless of whether they are based on the CT or DT model. Finally, an ICT procedure based on this frequency domain identification method is developed with the dynamic network. In the simulation, three types of defects in PCBA are demonstrated: short circuit, open circuit, and dynamic changes. Experiments show that for faulty full networks or subnetworks with only open circuits and dynamic changes, the ICT procedure accurately detects and diagnoses faults. This precision is maintained even with a single excitation, keeping the relative error for healthy components within $\pm 5\%$. In cases of full networks or subnetworks with short circuits at the edge nodes, the ICT procedure may not accurately estimate the values of components connected to the short circuits. However, it still effectively detects and isolates the short circuit using data and feature-level information. In general, the goal of the project has been achieved successfully.

B. Future work

There are some limitations that require further research. The first is that currently, we only focus on RLC circuits with passive components. There are also many active components in PCBA that we did not consider. An idea might be to use non-linear identification techniques to identify those components in PCBA [42]. The second is that, for the short-circuit cases, we only focus on the edge node, but they might happen everywhere in the PCBA in reality. More research is required for those tricky cases. The third is that the capacitors occur only in the interconnection with the ground nodes to satisfy identifiability. If the network identifiability condition can be further relaxed, the capacitances between the nodes without grounding can be estimated. The fourth is various computationally stable methods that can be employed to further improve numerically ill-conditioned frequency domain identification problems [43] [44]. Finally, all experiments are based on Matlab simulation; experiments in real setups in the real world can be performed to know about the real performance of this method.

VII. Appendix

A. The structure of Q_1 and θ_1 for criterion 1

1) Parameters of the structured network model θ_1

The polynomial matrix $A(p)$ has been defined in **Definition**. 1, for its (i,j) element defined as $a_{ij}(p) = \sum_{\ell=0}^{n_a} a_{ij,\ell} p^\ell$. The polynomial matrix $B(p)$ and the vector $C(p)$ also have a similar definition as $b_{ij}(p) = \sum_{\ell=0}^{n_b} b_{ij,\ell} p^\ell$ and $c_j(p) = \sum_{\ell=0}^{n_c} c_{j,\ell} p^\ell$. θ_1 collected all the parameters into a column vector as $\theta_1 = \begin{bmatrix} \theta_a^T & \theta_b^T & \theta_c^T \end{bmatrix}^T$. The vectors θ_a, θ_b and θ_c are given by,

$$\theta_a = \begin{bmatrix} \theta_{a1} \\ \theta_{a2} \\ \vdots \\ \theta_{aL} \end{bmatrix}, \quad \theta_{ai} = \begin{bmatrix} \theta_{a1j} \\ \theta_{a2j} \\ \vdots \\ \theta_{aLj} \end{bmatrix}, \quad \theta_{aij} = \begin{bmatrix} a_{ij,0} \\ a_{ij,1} \\ \vdots \\ a_{ij,n_a} \end{bmatrix}, \quad \text{with } i,j = 1,..,L \tag{73}$$

$$\theta_b = \begin{bmatrix} \theta_{b1} \\ \theta_{b2} \\ \vdots \\ \theta_{bL} \end{bmatrix}, \quad \theta_{bi} = \begin{bmatrix} \theta_{b1j} \\ \theta_{b2j} \\ \vdots \\ \theta_{bLj} \end{bmatrix}, \quad \theta_{bij} = \begin{bmatrix} b_{ij,0} \\ b_{ij,1} \\ \vdots \\ b_{ij,n_b} \end{bmatrix}, \quad \text{with } i = 1,..,L \text{ and } j = 1,..,K \tag{74}$$

$$\theta_c = \begin{bmatrix} \theta_{c1} \\ \theta_{c2} \\ \vdots \\ \theta_{cL} \end{bmatrix}, \quad \theta_{cj} = \begin{bmatrix} c_{j,0} \\ c_{j,1} \\ \vdots \\ c_{j,n_c} \end{bmatrix}, \text{ with } j = 1,..,L, \tag{75}$$

where, θ_a is a $L^2(n_a+1) \times 1$ vector, θ_b is a $LK(n_b+1) \times 1$ vector, and θ_c is a $L(n_c+1) \times 1$ vector.

*2) **The structure of the regressor** Q_1*

Recall the frequency weighting part $\mathfrak{W}$ for (37) as

$$\mathfrak{W} = \left[\hat{C}_W(k)A(\Omega_k,\theta)^{i-1}\right]^{-1} \tag{76}$$

which is a $L \times L$ matrix in each frequency point. Since the polynomial matrix $A(\Omega_k,\theta_a)^{i-1}$ for the last iteration is known, we can substitute the term Ω_k into it so that a $L \times L$ complex matrix is generated by taking the inverse of $A(\Omega_k,\theta_a)^{i-1}$ and multiplying another $L \times L$ complex matrix $\hat{C}_W^{-\frac{1}{2}}$ of each frequency sample data. Definition of Ω_{ka}, Ω_{kb}, and Ω_{kc} as a vector that collects the frequency variable of the matrices A, B, and C,

$$\begin{aligned}
\Omega_a &= \begin{bmatrix} 1 & \Omega_k & \Omega_k^2 & \cdots & \Omega_k^{n_a} \end{bmatrix}, \\
\Omega_b &= \begin{bmatrix} 1 & \Omega_k & \Omega_k^2 & \cdots & \Omega_k^{n_b} \end{bmatrix}, \\
\Omega_c &= \begin{bmatrix} 1 & \Omega_k & \Omega_k^2 & \cdots & \Omega_k^{n_c} \end{bmatrix}.
\end{aligned} \tag{77}$$

The polynomial matrix $A(p)$, $B(p)$ and $C(p)$ in frequency can be defined as $A(\Omega_k)$, $B(\Omega_k)$ and $C(\Omega_k)$, with theirs (i,j) element defined as,

$$\begin{aligned}
\tilde{a}_{ij}(\Omega_k) &= \sum_{\ell=0}^{n_a} a_{ij,\ell}\Omega_k^\ell, \\
\tilde{b}_{ij}(\Omega_k) &= \sum_{\ell=0}^{n_b} b_{ij,\ell}\Omega_k^\ell, \\
\tilde{c}_j(\Omega_k) &= \sum_{\ell=0}^{n_c} c_{j,\ell}\Omega_k^\ell.
\end{aligned} \tag{78}$$

Then, $M_1(k)$ in equation (37) for one frequency point can be derived as

$$\begin{aligned}
M_1 = &\begin{bmatrix} \mathfrak{W}_{A11}W_1 + \mathfrak{W}_{A12}W_2 + \cdots + \mathfrak{W}_{A1L}W_L \\ \mathfrak{W}_{A21}W_1 + \mathfrak{W}_{A22}W_2 + \cdots + \mathfrak{W}_{A2L}W_L \\ \vdots \\ \mathfrak{W}_{AL1}W_1 + \mathfrak{W}_{AL2}W_2 + \cdots + \mathfrak{W}_{ALL}W_L \end{bmatrix} \\
- &\begin{bmatrix} \mathfrak{W}_{B11}R_1 + \mathfrak{W}_{B12}R_2 + \cdots + \mathfrak{W}_{B1K}R_K \\ \mathfrak{W}_{B21}R_1 + \mathfrak{W}_{B22}R_2 + \cdots + \mathfrak{W}_{B2K}R_K \\ \vdots \\ \mathfrak{W}_{BL1}R_1 + \mathfrak{W}_{BL2}R_2 + \cdots + \mathfrak{W}_{BLK}R_K \end{bmatrix} \\
- &\begin{bmatrix} \mathfrak{W}_{C1} \\ \mathfrak{W}_{C2} \\ \vdots \\ \mathfrak{W}_{CL} \end{bmatrix},
\end{aligned} \tag{79}$$

where, $\mathfrak{W}_{Aij}$, $\mathfrak{W}_{Bij}$ and $\mathfrak{W}_{Cj}$ are given by,

$$\begin{aligned}
&\mathfrak{W}_{Aij} = \mathfrak{W}_{i1}\tilde{a}_{1j} + \mathfrak{W}_{i2}\tilde{a}_{2j} + \cdots + \mathfrak{W}_{iL}\tilde{a}_{Lj}, \\
&\text{with } i,j = 1,...,L,
\end{aligned} \tag{80}$$

$$\begin{aligned}
&\mathfrak{W}_{Bij} = \mathfrak{W}_{i1}\tilde{b}_{1j} + \mathfrak{W}_{i2}\tilde{b}_{2j} + \cdots + \mathfrak{W}_{iL}\tilde{b}_{Lj}, \\
&\text{with } i = 1,...,L \text{ and } j = 1,...,K,
\end{aligned} \tag{81}$$

$$\begin{aligned}
&\mathfrak{W}_{cj} = \mathfrak{W}_{j1}\tilde{c}_1 + \mathfrak{W}_{j2}\tilde{c}_2 + \cdots + \mathfrak{W}_{jL}\tilde{c}_L, \\
&\text{with } j = 1,...,L.
\end{aligned} \tag{82}$$

Extracting all the parameters defined as the column vector θ_1 in (73), (74) and (75) from $M_1(k)$, the regressor Q_1 is given by,

$$Q_1 = \begin{bmatrix} \Pi_{A1} & -\Pi_{B1} & -\Pi_{C1} \end{bmatrix}, \tag{83}$$

where, Π_{A1}, Π_{B1} and Π_{C1} are given as the following,

$$\Pi_{A1} = \begin{bmatrix} \mathfrak{W}_{W11} & \mathfrak{W}_{W12} & \cdots & \mathfrak{W}_{W1L} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{W}_{WL1} & \mathfrak{W}_{WL2} & \cdots & \mathfrak{W}_{WLL} \end{bmatrix}_{L\times L^2(n_a+1)},$$

$$\text{with } \mathfrak{W}_{Wij} = \begin{bmatrix} \mathfrak{W}_{ij}\Omega_a W_1 & \mathfrak{W}_{ij}\Omega_a W_2 & \cdots & \mathfrak{W}_{ij}\Omega_a W_L \end{bmatrix}, \tag{84}$$

$$\Pi_{B1} = \begin{bmatrix} \mathfrak{W}_{R11} & \mathfrak{W}_{R12} & \cdots & \mathfrak{W}_{R1K} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{W}_{RL1} & \mathfrak{W}_{R12} & \cdots & \mathfrak{W}_{RLK} \end{bmatrix}_{L\times LK(n_b+1)},$$

$$\text{with } \mathfrak{W}_{Rij} = \begin{bmatrix} , \mathfrak{W}_{ij}\Omega_b R_1 & \mathfrak{W}_{ij}\Omega_b R_2 & \cdots & \mathfrak{W}_{ij}\Omega_b R_K \end{bmatrix}, \tag{85}$$

and

$$\Pi_{C1} = \begin{bmatrix} \mathfrak{W}_{11}\Omega_c & \mathfrak{W}_{12}\Omega_c & \cdots & \mathfrak{W}_{1L}\Omega_c \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{W}_{L1}\Omega_c & \mathfrak{W}_{L2}\Omega_c & \cdots & \mathfrak{W}_{LL}\Omega_c \end{bmatrix}_{L\times(n_c+1)}. \tag{86}$$

By using the Kronecker product, they can be derived as

$$\begin{aligned}
\Pi_{A1} &= \mathfrak{W} \otimes \begin{bmatrix} \Omega_a W_1 & \Omega_a W_2 & \cdots & \Omega_a W_L \end{bmatrix}, \\
\Pi_{B1} &= \mathfrak{W} \otimes \begin{bmatrix} \Omega_b R_1 & \Omega_b R_2 & \cdots & \Omega_b R_K \end{bmatrix}, \\
\Pi_{C1} &= \mathfrak{W} \otimes \Omega_c.
\end{aligned} \tag{87}$$

The dimensions of Q_1 for each frequency point is $L \times [L^2 \times (n_a+1) + LK(n_b+1) + L(n_c+1)]$, and for the number of F frequency data, Q_1 is collected in $L \times [L^2(n_a+1) + LK(n_b+1) + L(n_c+1)] \times F$.

B. The structure of Q_2 and θ_2 for criterion 2

*1) **Parameters of the structured network model** θ_2*

For this case, since the transient term was assumed that has been sufficiently removed in the non-parametric identification, the parametric modal will not include the transient term model parameters. Therefore, θ_2 is given by

$$\theta_2 = \begin{bmatrix} \theta_a^T & \theta_b^T \end{bmatrix}^T, \tag{88}$$

where, the vectors θ_a and θ_b have the same structure as the previous definition (73) and (74), the dimension of them are also the same as given before.

*2) **The structure of the regressor** Q_2*

Recall the definition of Ω_{ka}, Ω_{kb} in (77) and the definition of the element (i,j) for polynomial matrix $A(p)$,$B(p)$ for frequency in (78), $M_2(k)$ can be derived as

$$M_2 = \begin{bmatrix} \mathfrak{W}_{A11}\mathfrak{G}_{R1} + \mathfrak{W}_{A12}\mathfrak{G}_{R1} + \cdots + \mathfrak{W}_{A1L}\mathfrak{G}_{RL} \\ \mathfrak{W}_{A21}\mathfrak{G}_{R1} + \mathfrak{W}_{A22}\mathfrak{G}_{R2} + \cdots + \mathfrak{W}_{A2L}\mathfrak{G}_{RL} \\ \vdots \\ \mathfrak{W}_{AL1}\mathfrak{G}_{R1} + \mathfrak{W}_{AL2}\mathfrak{G}_{R2} + \cdots + \mathfrak{W}_{ALL}\mathfrak{G}_{RL} \end{bmatrix} - \begin{bmatrix} \mathfrak{W}_{B11}R_1 + \mathfrak{W}_{B12}R_2 + \cdots + \mathfrak{W}_{B1K}R_K \\ \mathfrak{W}_{B21}R_1 + \mathfrak{W}_{B22}R_2 + \cdots + \mathfrak{W}_{B2K}R_K \\ \vdots \\ \mathfrak{W}_{BL1}R_1 + \mathfrak{W}_{BL2}R_2 + \cdots + \mathfrak{W}_{BLK}R_K \end{bmatrix}, \tag{89}$$

where, $\mathfrak{W}_{Aij}$, $\mathfrak{W}_{Bij}$ are keeping the same definition in (80) and (81), while $\mathfrak{G}_{Rj}$ is given by

$$\begin{aligned} &\mathfrak{G}_{Rj} = (\hat{G}_{j1}R_1 + \hat{G}_{j2}R_2 + \cdots + \hat{G}_{jK}R_K), \\ &\text{with } j = 1, ..., L \end{aligned} \tag{90}$$

Extracting all the parameters defined as the column vector θ_2 in (73), (74) from $M_2(k)$, the regressor Q_2 is given by,

$$Q_2 = \begin{bmatrix} \Pi_{A2} & -\Pi_{B2} \end{bmatrix}, \tag{91}$$

where Π_{A2} and Π_{B2} are given by,

$$\Pi_{A2} = \begin{bmatrix} \mathfrak{W}_{G11} & \mathfrak{W}_{G11} & \cdots & \mathfrak{W}_{G1L} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{W}_{GL1} & \mathfrak{W}_{GL2} & \cdots & \mathfrak{W}_{GLL} \end{bmatrix}_{L \times L^2(n_a+1)}$$
$$\text{with } \mathfrak{W}_{Gij} = \begin{bmatrix} \mathfrak{W}_{ij}\Omega_a\mathfrak{G}_{R1} & \mathfrak{W}_{ij}\Omega_a\mathfrak{G}_{R2} & \cdots & \mathfrak{W}_{ij}\Omega_a\mathfrak{G}_{RL} \end{bmatrix} \tag{92}$$

$$\Pi_{B2} = \begin{bmatrix} \mathfrak{W}_{R11} & \mathfrak{W}_{R12} & \cdots & \mathfrak{W}_{R1K} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{W}_{RL1} & \mathfrak{W}_{R12} & \cdots & \mathfrak{W}_{RLK} \end{bmatrix}_{L \times LK(n_b+1)}$$
$$\text{with } \mathfrak{W}_{Rij} = \begin{bmatrix} \mathfrak{W}_{ij}\Omega_b R_1 & \mathfrak{W}_{ij}\Omega_b R_2 & \cdots & \mathfrak{W}_{ij}\Omega_b R_K \end{bmatrix}. \tag{93}$$

By using the Kronecker product, they can be derived as

$$\begin{aligned} \Pi_{A2} &= \mathfrak{W} \otimes \begin{bmatrix} \Omega_a\mathfrak{G}_{R1} & \Omega_a\mathfrak{G}_{R2} & \cdots & \Omega_a\mathfrak{G}_{RL} \end{bmatrix}, \\ \Pi_{B2} &= \mathfrak{W} \otimes \begin{bmatrix} \Omega_b R_1 & \Omega_b R_2 & \cdots & \Omega_b R_K \end{bmatrix}. \end{aligned} \tag{94}$$

The dimensions of Q_2 for each frequency point is $L \times [L^2(n_a + 1) + LK(n_b + 1)]$, and for the number of F frequency data, Q_2 can be collected in $L \times [L^2(n_a + 1) + LK(n_b + 1))] \times F$.

C. *Lagrangian multiplier optimization constraints*

For the diffusively coupled network model, the A matrix is the symmetric polynomial matrix, which means that for all $i \neq j$ we have $a_{ij} - a_{ji} = 0$. To build the constraint for this, Γ_a and υ_a are given by

$$\Gamma_a = \begin{bmatrix} \Gamma_{a,1} \\ \Gamma_{a,2} \\ \vdots \\ \Gamma_{a,(L-1)} \end{bmatrix}, \quad \upsilon_a = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{\frac{L(L-1)}{2}(n_a+1)\times 1}, \tag{95}$$

where $\Gamma_{a,i}$ with $i = 1, ..., L-1$ is given by

$$\begin{aligned} &\Gamma_{a,i} = \begin{bmatrix} \mathrm{O}_{d_1 \times d_2} & \mathrm{I}_{d_3 \times d_4} & \gamma_i \end{bmatrix}, \\ &\text{with } \gamma_i = \begin{bmatrix} \gamma_{1,i} & \gamma_{2,i} & \cdots & \gamma_{(L-i),i} \end{bmatrix}, \end{aligned} \tag{96}$$

where O is the zero matrix with the dimension $d_1 \times d_2 = (L - i)(n_a + 1) \times (n_a + 1)(L(i - 1) + i)$ and I is the identity matrix with the dimension $d_3 \times d_4 = (L - i)(n_a + 1) \times (L - i)(n_a + 1)$. Notice that γ_i is the matrix with the change dimension for the change i, the dimension is defined as $(L - i)(n_a + 1) \times (n_a + 1)L$ so that this matrix can be considered as a matrix consisting of $(L - i) \times L$ blocks and each block is a submatrix $(n_a + 1) \times (n_a + 1)$. The $[(L - i), i]$-th block of this γ_i matrix is defined as the identity submatrix with the negative (-) sign. The dimension of Γ_a is with the dimension $\frac{L(L-1)}{2}(n_a + 1) \times L^2(n_a + 1)$. By restricting $\Gamma_a\theta_a = \upsilon_a$, the symmetric structure of the polynomial matrix A is guaranteed.

Next, we set the constraint for the partially known dynamics or known parameters of the model. Here, we assume that all the dynamics of the input signals are known, which means that all the parameters of the polynomial matrix B are known. Therefore, the constraint for B as is given by

$$\Gamma_b = \begin{bmatrix} \mathrm{I}_{n_b+1} & & \\ & \ddots & \\ & & \mathrm{I}_{n_b+1} \end{bmatrix}, \tag{97}$$

with the dimension $LK(n_b+1) \times LK(n_b+1)$ and υ_b collects all the known parameters of B in the same structure of θ_b in (74). Moreover, if some component dynamics is already known, we can also set the constraint with the known real parameters in the A matrix in the same way. Finally, the constraints Γ and υ collect all the above constraints, which are given by,

$$\Gamma = \begin{bmatrix} \Gamma_a & \mathrm{O}_{d_5 \times d_6} & \mathrm{O}_{d_9 \times d_{10}} \\ \mathrm{O}_{d_7 \times d_8} & \Gamma_b & \mathrm{O}_{d_{11} \times d_{12}} \end{bmatrix}, \upsilon = \begin{bmatrix} \upsilon_a \\ \upsilon_b \end{bmatrix}, \tag{98}$$

where, O are the zero matrices with the dimension $d_5 \times d_6 = \frac{L(L-1)}{2}(n_a + 1) \times LK(n_b + 1)$, $d_7 \times d_8 = LK(n_b + 1) \times L^2(n_a + 1)$, $d_9 \times d_{10} = \frac{L(L-1)}{2}(n_a + 1) \times L(n_c + 1)$, and $d_{11} \times d_{12} = LK(n_b + 1) \times L(n_c + 1)$.

Remark 18. Here, if there is no transient term estimation, define $n_c = -1$.

Algorithm 1 Γ_a Matrix Construction Algorithm

Initialize FM1 and FM2 matrices
for $k = 1$ **to** $L * (L-1)/2$ **do**
  $item \leftarrow 0$
  $flag \leftarrow 0$
  for $i = 1$ **to** L **do**
    for $j = 1$ **to** L **do**
      $item \leftarrow item + 1$
      if $i < j$ **and** $FM1(i,j) = 0$ **and** $flag = 0$ **then**
        $\Gamma_a[k, item] \leftarrow \text{Identity}_{(n_a+1)}$
        $FM1(i,j) \leftarrow 1$
        $FM1(j,i) \leftarrow 1$
        $flag \leftarrow 1$
      else if $i > j$ **and** $FM1(j,i) = 1$ **and** $FM2(i,j) \neq 0$ **then**
        $\Gamma_a[k, item] \leftarrow$ - $\text{Identity}_{(n_a+1)}$
        $FM2(i,j) \leftarrow 0$
      else
        $\Gamma_a[k, item] \leftarrow \text{Zero}_{(n_a+1)}$
      end if
    end for
    if $i == L$ **then**
      $\Gamma\_a[k, item+1] \leftarrow \text{Zero}_{(n_a+1)}$
    end if
  end for
end for

D. Code checking example

This example is to double check the constraint selection matrix Γ_a. The idea of this example is to select the $A_{ij,0}, ...A_{ij,n_a}$ by using identity matrix Identity_{n_a+1} and minus $A_{ji,0}, ...A_{ji,n_a}$ by using negative identity matrix - $\text{Identity}_{(n_a+1)}$, FM1 and FM2 are matrices used as flags to track whether a pair (i,j) or its inverse (j,i) has been processed. By comparing the difference of matrices in two-way generating, the code implementation could be checked

E. The structure of Q_3 and θ_3 for estimating the target subnetwork

1) The parameters vector θ_3

$\theta_3 = \begin{bmatrix}\theta_{Ajj}^T & \theta_{Ajd}^T & \theta_{Bj}^T\end{bmatrix}$ The vectors θ_{Ajj}, θ_{Ajd} and θ_{Bj} are given by,

$$\theta_{Ajj} = \begin{bmatrix}\theta_{a1}\\ \theta_{a2}\\ \vdots \\ \theta_{aL_J}\end{bmatrix}, \quad \theta_{ai} = \begin{bmatrix}\theta_{a1j}\\ \theta_{a2j}\\ \vdots \\ \theta_{aL_Jj}\end{bmatrix}, \quad \theta_{aij} = \begin{bmatrix}a_{ij,0}\\ a_{ij,1}\\ \vdots \\ a_{ij,n_a}\end{bmatrix},$$
with $i, j = 1, .., L_J$ (99)

$$\theta_{Ajd} = \begin{bmatrix}\theta_{a1}\\ \theta_{a2}\\ \vdots \\ \theta_{aL_J}\end{bmatrix}, \quad \theta_{ai} = \begin{bmatrix}\theta_{a1j}\\ \theta_{a2j}\\ \vdots \\ \theta_{aL_Jj}\end{bmatrix}, \quad \theta_{aij} = \begin{bmatrix}a_{ij,0}\\ a_{ij,1}\\ \vdots \\ a_{ij,n_a}\end{bmatrix},$$
with $i = 1, .., L_J$ and $j = L_J + 1, .., L_S$ (100)

$$\theta_{Bj} = \begin{bmatrix}\theta_{b1}\\ \theta_{b2}\\ \vdots \\ \theta_{bL_J}\end{bmatrix}, \quad \theta_{bi} = \begin{bmatrix}\theta_{b1j}\\ \theta_{b2j}\\ \vdots \\ \theta_{bL_Jj}\end{bmatrix}, \quad \theta_{bij} = \begin{bmatrix}b_{ij,0}\\ b_{ij,1}\\ \vdots \\ b_{ij,n_b}\end{bmatrix},$$
with $i = 1, .., L_J$ and $j = 1, .., K$ (101)

where, $L_J \times L_J$ is the dimension of the target subnetwork matrix A_{JJ}, $L_J \times L_S$ is dimension of the immersed network $A_{im}(p)$, θ_{Ajj} is a vector $L_J^2(n_a+1) \times 1$, θ_{Ajd} is a vector $L_J L_D(n_a+1) \times 1$ with $L_D = L_S - L_J$ and θ_{Bj} is a vector $L_J K(n_b+1) \times 1$. Collecting the parameters θ_{Ajj} and θ_{Ajd} in the matrix A independently can make the construction of the regressor matrix more convenient.

2) The structure of the regressor Q_3

Recall the definition of Ω_{ka}, Ω_{kb} in (77) and the definition of the element (i,j) for the polynomial matrix $A(p)$,$B(p)$ for frequency in (78) but separated for A_{JJ} and A_{JD} as, element defined as,

$$\begin{aligned}\tilde{a}_{J_{ij}}(\Omega_k) &= \sum_{\ell=0}^{n_a} a_{JJ_{ij},\ell}\Omega_k^\ell,\\ \tilde{a}_{D_{ij}}(\Omega_k) &= \sum_{\ell=0}^{n_a} a_{JD_{ij},\ell}\Omega_k^\ell,\\ \tilde{b}_{ij}(\Omega_k) &= \sum_{\ell=0}^{n_b} b_{J_{ij},\ell}\Omega_k^\ell.\end{aligned} \tag{102}$$

Define the data vectors for the frequency sample Ω_{kaim}, Ω_{kbim} for the immersed network, which maintains the same structure as (77) but with different order as $n_{a_{im}}$ and $n_{b_{im}}$, respectively. The known immersed polynomial matrix $\hat{A}_{im}(\hat{\eta})$ and $\hat{B}_{im}(\hat{\eta})$ in frequency can be defined as $\hat{A}_{im}(\Omega_k, \hat{\eta})$ and $\hat{B}_{im}(\Omega_k, \hat{\eta})$, with theirs (i,j) element defined as,

$$\begin{aligned}\hat{a}_{IMJ_{ij}}(\Omega_k) &= \sum_{\ell=0}^{n_{a_{im}}} a_{IMJJ_{ij},\ell}\Omega_k^\ell,\\ \hat{a}_{IMD_{ij}}(\Omega_k) &= \sum_{\ell=0}^{n_{a_{im}}} a_{IMJD_{ij},\ell}\Omega_k^\ell,\\ \hat{b}_{IMJ_{ij}}(\Omega_k) &= \sum_{\ell=0}^{n_{b_{im}}} b_{IMJ_{ij},\ell}\Omega_k^\ell,\end{aligned} \tag{103}$$

where $\hat{a}_{IMJ_{ij}}(\Omega_k)$ and $\hat{a}_{IMD_{ij}}(\Omega_k)$ are the estimation frequency response data with each frequency point for each element in the estimated $\hat{A}_{JJ}$ and $\hat{A}_{JD}$, respectively. The first equation in (67) can be expanded as.

$$M_3 = \begin{bmatrix} \mathfrak{J}_{a11} & \mathfrak{J}_{a12} & \cdots & \mathfrak{J}_{a1L_D} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{J}_{aL_J1} & \mathfrak{J}_{aL_J2} & \cdots & \mathfrak{J}_{aL_JL_D} \end{bmatrix} - \begin{bmatrix} \mathfrak{D}_{a11} & \mathfrak{D}_{a11} & \cdots & \mathfrak{D}_{a1L_D} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{D}_{aL_J1} & \mathfrak{D}_{aL_J2} & \cdots & \mathfrak{D}_{aL_JL_D} \end{bmatrix}, \tag{104}$$

where, $\mathfrak{J}_{aij}$ and $\mathfrak{D}_{aij}$ are given by

$$\begin{aligned} &\mathfrak{J}_{aij} = \hat{a}_{IMD_{1j}}\tilde{a}_{J_{i1}} + \hat{a}_{IMD_{2j}}\tilde{a}_{J_{i2}} + \cdots + \hat{a}_{IMD_{L_Jj}}\tilde{a}_{J_{iL_J}}, \\ &\mathfrak{D}_{aij} = \hat{a}_{IMJ_{i1}}\tilde{a}_{D_{1j}} + \hat{a}_{IMJ_{i2}}\tilde{a}_{D_{2j}} + \cdots + \hat{a}_{IMJ_{iL_J}}\tilde{a}_{D_{L_Jj}}, \\ &\text{with } i = 1, ..., L_J \text{ and } j = 1, ..., L_D. \end{aligned} \tag{105}$$

The second equation in (67) can be expanded as

$$M_4 = \begin{bmatrix} \mathfrak{J}_{b11} & \mathfrak{J}_{b12} & \cdots & \mathfrak{J}_{b1K} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{J}_{bL_J1} & \mathfrak{J}_{bL_J2} & \cdots & \mathfrak{J}_{bL_JK} \end{bmatrix} - \begin{bmatrix} \mathfrak{B}_{b11} & \mathfrak{B}_{b11} & \cdots & \mathfrak{B}_{b1K} \\ \vdots & \vdots & \vdots & \vdots \\ \mathfrak{B}_{bL_J1} & \mathfrak{B}_{bL_J2} & \cdots & \mathfrak{B}_{bL_JK} \end{bmatrix}, \tag{106}$$

where, $\mathfrak{J}_{bij}$ and $\mathfrak{B}_{bij}$ are given by

$$\begin{aligned} &\mathfrak{J}_{bij} = \hat{b}_{IMJ_{1j}}\tilde{a}_{J_{i1}} + \hat{b}_{IMJ_{2j}}\tilde{a}_{J_{i2}} + \cdots + \hat{b}_{IMJ_{L_Jj}}\tilde{a}_{J_{iL_J}}, \\ &\mathfrak{B}_{bij} = \hat{a}_{IMJ_{i1}}\tilde{b}_{J_{1j}} + \hat{a}_{IMJ_{i2}}\tilde{b}_{J_{2j}} + \cdots + \hat{a}_{IMJ_{iL_J}}\tilde{b}_{J_{L_Jj}}, \\ &\text{with } i = 1, ..., L_J \text{ and } j = 1, ..., K. \end{aligned} \tag{107}$$

Extracting all the parameters defined as the column vector θ_3 in (100), (100) and (101) from M_3 and M_4, the regressor Q_3 is given by

$$Q_3 = \begin{bmatrix} \Pi_{A3} & \Pi_{B3} \end{bmatrix}^T, \tag{108}$$

with

$$\begin{aligned} \Pi_{A3} &= \begin{bmatrix} \Pi_{A31} & -\Pi_{A32} \end{bmatrix}, \\ \Pi_{A31} &= \mathrm{I}_{L_J} \otimes (\hat{A}_{JD}(\Omega_k, \hat{\eta})^T \otimes \Omega_a), \\ \Pi_{A32} &= \hat{A}_{JJ}(\Omega_k, \hat{\eta}) \otimes (\mathrm{I}_{L_D} \otimes \Omega_a), \end{aligned} \tag{109}$$

and

$$\begin{aligned} \Pi_{B3} &= \begin{bmatrix} \Pi_{B31} & -\Pi_{B31} \end{bmatrix}, \\ \Pi_{B31} &= \mathrm{I}_{L_J} \otimes (\hat{B}_{J}(\Omega_k, \hat{\eta})^T \otimes \Omega_a), \\ \Pi_{B32} &= \hat{A}_{JJ}(\Omega_k, \hat{\eta}) \otimes (\mathrm{I}_{K} \otimes \Omega_b). \end{aligned} \tag{110}$$

The dimensions of Q_3 for each frequency point is $[L_J(L_D + K)] \times [L_JL_S(n_a + 1) + L_JK(n_b + 1)]$, and for the number of F frequency data, Q_3 can be collected in $[L_J(L_D + K)] \times [L_JL_S(n_a + 1) + L_JK(n_b + 1)] \times F$.

F. FRF estimation for a faulty 7 node network

The 7-node faulty network model example with only open circuit in L_{34} is shown here. The full excited and full measurement setup with the same types of excitation signals and noise signals as in Example 2 is given. The frequency band is set between 500 Hz and 6000 Hz. The FRF estimation of this continuous-time seven-node defect model with the true FRF of the healthy model is shown in Fig. 23.

ACKNOWLEDGMENT

Firstly, my deepest gratitude goes to my supervisors. A special thanks to Paul for providing me with the opportunity to pursue research and for your invaluable guidance. I would also like to thank Lizan for your patience in answering my questions. Your insightful questions and feedback have been instrumental in my development. Secondly, a big thank you to my MSc colleagues. Our engaging discussions and mutual support have been invaluable. I also want to thank Zhiyong for consistently being accessible to help students and for providing heartfelt suggestions. I must also thank Yunxian for your continued support and love. Your encouragement has been a pillar of our shared growth from Bachelor's to Master's degrees. Lastly, I would like to express my deep thanks to Mom and Dad for your support during my studies abroad. You have always been my greatest support. Thanks to all those who have helped me throughout this journey.

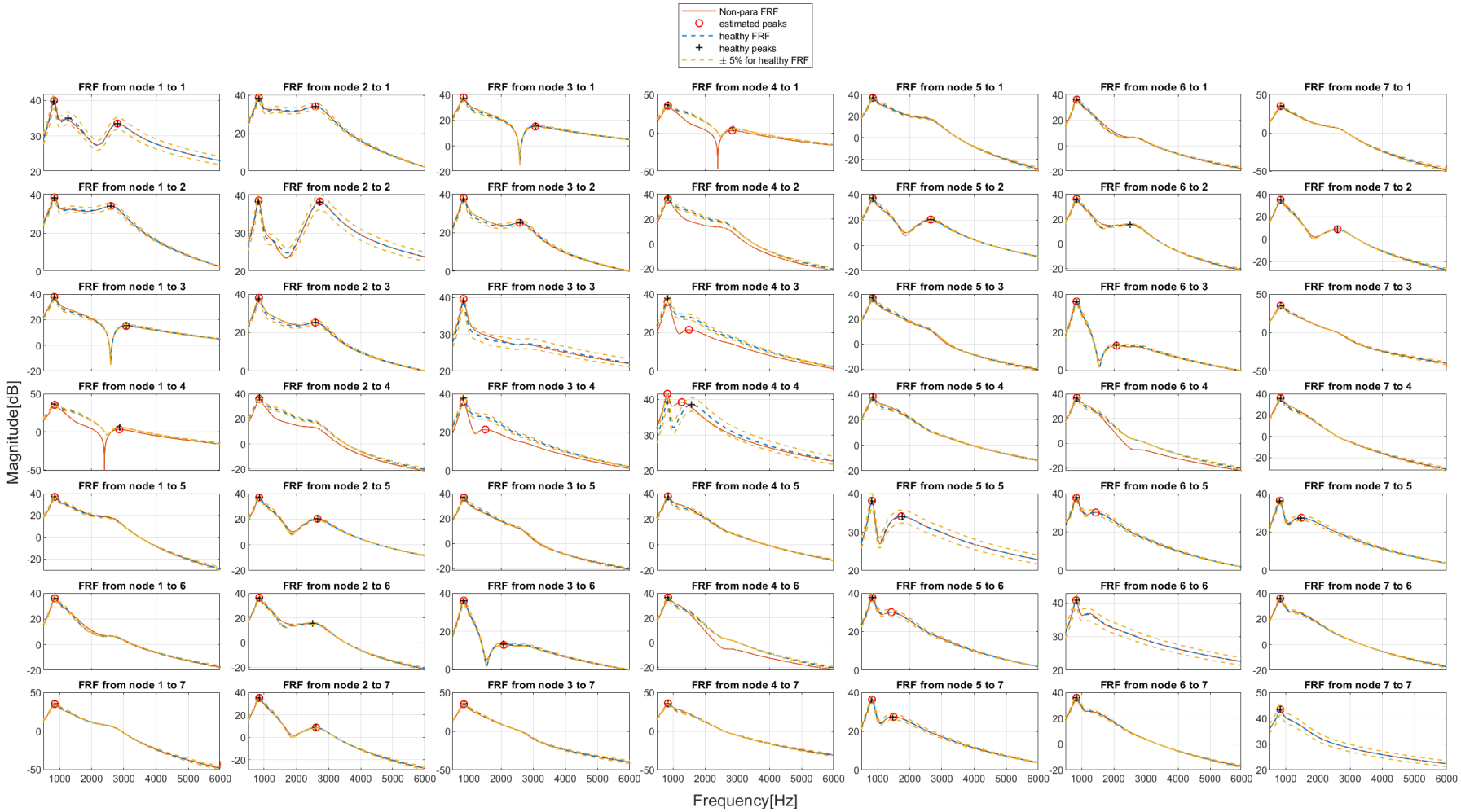


Figure 23: FRF estimation of defect model (L_{34} open circuit) with the true FRF of the healthy model (CT model)

REFERENCES

[1] P. M. A. Vitoriano, T. G. Amaral, and O. P. Dias, "Automatic optical inspection for surface mounting devices with ipc-a-610d compliance," in *2011 International Conference on Power Engineering, Energy and Electrical Drives*, pp. 1–7, 2011.

[2] P. Express, "X-ray inspection: What is it and how does it work?," 2021. Available at https://www.protoexpress.com/kb/x-ray-inspection/.

[3] D. Miljković, "Fault detection methods: A literature survey," in *2011 Proceedings of the 34th International Convention MIPRO*, pp. 750–755, 2011.

[4] T. Söderström and P. Stoica, *System identification*. Prentice-Hall International, 1989.

[5] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Electrical networks and algebraic graph theory: Models, properties, and applications," *Proceedings of the IEEE*, vol. 106, no. 5, pp. 977–1005, 2018.

[6] L. Ljung, "System identification," in *Signal analysis and prediction*, pp. 163–173, Springer, 1998.

[7] X. Cheng and J. Scherpen, "Model reduction methods for complex network systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, pp. 425–453, 2021.

[8] M. Friswell, S. Garvey, and J. Penny, "Extracting second-order systems from state-space representations," *AIAA journal*, vol. 37, no. 1, pp. 132–135, 1999.

[9] H. Luş, M. De Angelis, R. Betti, and R. W. Longman, "Constructing second-order models of mechanical systems from identified state space realizations. part i: theoretical discussions," *Journal of Engineering Mechanics*, vol. 129, no. 5, pp. 477–488, 2003.

[10] M. De Angelis, H. Lus, R. Betti, and R. W. Longman, "Extracting physical parameters of mechanical models from identified state-space representations," *J. Appl. Mech.*, vol. 69, no. 5, pp. 617–625, 2002.

[11] E. M. M. Kivits and P. M. J. Van den Hof, "A dynamic network approach to identification of physical systems," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, pp. 4533–4538, 2019.

[12] E. M. M. Kivits and P. M. J. Van den Hof, "Identification of diffusively coupled linear networks through structured polynomial models," *IEEE Transactions on Automatic Control*, vol. 68, no. 6, pp. 3513–3528, 2023.

[13] E. M. M. Kivits and P. M. J. Van den Hof, "Local identification in diffusively coupled linear networks," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, pp. 874–879, 2022.

[14] J. Meijer, "Improvements for in-circuit testing using rlc network identification," Master's thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 2021.

[15] R. A. González, C. R. Rojas, S. Pan, and J. S. Welsh, "On the relation between discrete and continuous-time refined instrumental variable methods," *IEEE Control Systems Letters*, vol. 7, pp. 2233–2238, 2023.

[16] H. Garnier, M. Gilson, H. Muller, and F. Chen, "A new graphical user interface for the CONTSID toolbox for matlab," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 397–402.

[17] A. G. Dankers, P. M. J. Van den Hof, and X. Bombois, "Direct and indirect continuous-time identification in dynamic networks," in *53rd IEEE Conference on Decision and Control*, pp. 3334–3339, 2014.

[18] H. Garnier, L. Wang, and P. C. Young, *Direct identification of continuous-time models from sampled data: Issues, basic solutions and relevance*. Springer, 2008.

[19] M. Gilson, H. Garnier, P. C. Young, and P. M. J. Van den Hof, "Instrumental variable methods for closed-loop continuous-time model identification," *Identification of continuous-time models from sampled data*, pp. 133–160, 2008.

[20] A. Dankers, P. M. J. Van den Hof, X. Bombois, and P. S. Heuberger, "Errors-in-variables identification in dynamic networks by an instrumental variable approach," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 2335–2340, 2014.

[21] K. R. Ramaswamy, P. Z. Csurcsia, J. Schoukens, and P. M. J. Van den Hof, "A frequency domain approach for local module identification in dynamic networks," *Automatica*, vol. 142, p. 110370, 2022.

[22] D. Liang, "Continuous-time identification of diffusively coupled networks." Internship Report, 2023.

[23] P. Express, "Basic components overview," 2021. Available at https://www.protoexpress.com/kb/basic-components-overview/.

[24] Seeedstudio, "13 common pcb soldering problems to avoid," 2021. Available at https://www.seeedstudio.com/blog/2021/06/18/13-common-pcb-soldering-problems-to-avoid/.

[25] M. Mesbahi and M. Egerstedt, "Graph theoretic methods in multiagent networks," in *Graph Theoretic Methods in Multiagent Networks*, Princeton University Press, 2010.

[26] R. Pintelon, J. Schoukens, and H. Chen, "On the basic assumptions in the identification of continuous time systems," *IFAC Proceedings Volumes*, vol. 27, no. 8, pp. 1105–1114, 1994.

[27] R. Pintelon and J. Schoukens, *System identification: a frequency domain approach*. John Wiley & Sons, 2012.

[28] R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé, "Estimation of nonparametric noise and frf models for multivariable systems—part i: Theory," *Mechanical Systems and Signal Processing*, vol. 24, no. 3, pp. 573–595, 2010.

[29] R. Pintelon, J. Schoukens, and P. Guillaume, "Continuous-time noise modeling from sampled data," *IEEE Transactions on Instrumentation and Measurement*, vol. 55, no. 6, pp. 2253–2258, 2006.

[30] J. Schoukens, Y. Rolain, and R. Pintelon, "Analysis of windowing/leakage effects in frequency response function measurements," *Automatica*, vol. 42, no. 1, pp. 27–38, 2006.

[31] F. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, 1978.

[32] I. Kollár, "Frequency domain system identification toolbox for matlab," 2004-2019.

[33] E. M. M. Kivits and P. M. J. Van den Hof, "Identifiability of diffusively coupled linear networks with partial instrumentation," *IFAC-PapersOnLine*, vol. 52, no. 2, p. 2395–2400, 2023. 22nd World Congress of the International Federation of Automatic Control (IFAC 2023 World Congress) ; Conference date: 09-07-2023 Through 14-07-2023.

[34] G. Strang, *Linear Algebra and Its Applications*. Thomson, Brooks/Cole, 2006.

[35] C. Sanathanan and J. Koerner, "Transfer function synthesis as a ratio of two complex polynomials," *IEEE transactions on automatic control*, vol. 8, no. 1, pp. 56–58, 1963.

[36] R. Pintelon, J. Schoukens, G. Vandersteen, and K. Barbé, "Estimation of nonparametric noise and frf models for multivariable systems—part ii: Extensions, applications," *Mechanical Systems and Signal Processing*, vol. 24, no. 3, pp. 596–616, 2010.

[37] R. de Callafon, D. de Roover, and P. M. J. Van den Hof, "Multivariable least squares frequency domain identification using polynomial matrix fraction descriptions," in *Proceedings of 35th IEEE Conference on Decision and Control*, vol. 2, pp. 2030–2035 vol.2, 1996.

[38] E. K. Chong and S. H. Zak, *An introduction to optimization*, vol. 75. John Wiley & Sons, 2013.

[39] R. Pintelon and I. Kollar, "On the frequency scaling in continuous-time modeling," *IEEE Transactions on Instrumentation and Measurement*, vol. 54, no. 1, pp. 318–321, 2005.

[40] F. Dörfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.

[41] A. Dankers, P. M. J. Van den Hof, X. Bombois, and P. S. C. Heuberger, "Identification of dynamic models in complex networks with prediction error methods: Predictor input selection," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 937–952, 2016.

[42] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented road map," *IEEE Control Systems Magazine*, vol. 39, no. 6, pp. 28–99, 2019.

[43] R. S. Blom and P. M. J. Van den Hof, "Multivariable frequency domain identification using iv-based linear regression," in *49th IEEE Conference on Decision and Control (CDC)*, pp. 1148–1153, 2010.

[44] R. van Herpen, T. Oomen, and M. Steinbuch, "Optimally conditioned instrumental variable approach for frequency-domain system identification," *Automatica*, vol. 50, no. 9, pp. 2281–2293, 2014.