Submitted to the 20th IFAC Symposium on System Identification, 17-19 July 2024, Boston, MA, USA. 10 December 2023.

Fault detection and diagnosis using the dynamic network framework ⋆

Yibo Shi, **Stefanie J. M. Fonken** and **Paul M. J. Van den Hof**

Control Systems Group, Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands. (e-mail:syb.hn.china@gmail.com; s.j.m.fonken@tue.nl; p.m.j.vandenhof@tue.nl).

Abstract: A local model-based method for fault detection and diagnosis (FDD) in large-scale interconnected network systems is introduced, using models in a dynamic network framework. To this end, model validation methods are developed for validating single modules in a dynamic network, which are generalized from the classical auto- and cross-correlation tests for open- and closed-loop systems. Invalidation of the model can indicate the detection of a fault in the system. A fault diagnosis algorithm is developed that includes fault isolation and optimal placement of external excitation signals. Numerical illustrations demonstrate the method's capability to detect a fault in a local module and isolate it within the entire network system.

Keywords: Fault, interconnected system, dynamic network, model validation, correlation test.

1. INTRODUCTION

Modern large-scale interconnected network systems, such as power grids, distributed control systems, and communication networks, are essential in various areas of daily life. In recent years, there has been a growing interest in developing data-driven modeling for those network systems within the field of system identification, employing the dynamic network framework, see e.g. Gonçalves and Warnick (2008); Van den Hof et al. (2013); Gevers et al. (2019). A dynamic network model describes the input–output behavior between each pair of observed signals in a network system, as well as its spatial structure, referred to as the network topology. This dynamic network framework provides a comprehensive perspective for understanding and analyzing network systems and facilitates the development of advanced fault detection and diagnosis (FDD) techniques. This research focuses on the development of FDD methods for large-scale interconnected network systems using dynamic network models. In addition, the availability of sophisticated sensors, actuators, and processors makes local FDD more attractive, where the system is divided into subsystems and local detection and diagnosis can be developed for each subsystem (Khorasgani et al. (2015)).

Dowdeswell et al. (2020) emphasizes the crucial importance of reliable models when developing model-based FDD methods for network systems. This requirement can be met through dynamic network identification, which has yielded several methods for consistgently estimating either a full network or a local subnetwork. (Materassi and Salapaka (2012); Van den Hof et al. (2013); Dankers et al. (2016); Weerts et al. (2018); Gevers et al. (2019); Ramaswamy et al. (2022); Fonken et al. (2022)). Furthermore, model-based FDD is based on the concept of model invalidation: if the collected data invalidates an existing accurate model, it can indicate potential faults, possibly with information for further diagnosis (Gertler (1998)). Since the specific challenge of validating an estimated dynamic network model has not been addressed, our study poses the following research question: How can we perform local model-based FDD in a connected network system while using its model within the dynamic network framework?

To tackle this research question, we will first develop a framework for model validation of local modules in a dynamic network. In our approach, we will decompose the network into multi-input-single-output (MISO) subnetworks for which model validation tools based on residual analysis will be developed. The typical situation for dynamic networks is that the validation tests typically provide information on a set of modules so that the tests will need to be combined appropriately to arrive at statements on single modules.

After presenting the network setup in Section 2, we will present basic correlation tests to be used as model (in)validation tests in the dynamic network framework. Sections 4 and 5 describe how to use the model validation tests for FDD respectively. Numerical illustrations of the developed FDD procedure with a 3-node network model are collected in Section 6.

2. NETWORK SETUP

2.1 MISO subnetwork setup

Following the setup as in Van den Hof et al. (2013), a dynamic network model is built up of L node signals $w_j(t),\ j \in \mathcal{L}$ with $\mathcal{L} = [1,\ L]$ the index set of all node signals. The network can be written in a MISO structure:

⋆ This project has received funding from the European Research Council (ERC), Advanced Research Grant SYSDYNET, under the European Union's Horizon 2020 research and innovation programme (grant agreement No 694504).

$$w_j(t) = \sum_{k\in\mathcal{N}_j} G^0_{jk}(q)w_k(t) + r_j(t) + v_j(t), \qquad (1)$$

where q^{-1} is the delay operator, i.e. $q^{-1}u(t) = u(t-1)$;

- $G^0_{jk}(z)$ is a strictly proper rational transfer function, representing a direct causal connection from w_k to w_j;
- $\mathcal{N}_j$ is the set of indices of node signals w_k, $k \neq j$, for which $G^0_{jk} \neq 0$, referred to as the in-neighbors of node w_j;
- v_j is a process noise, where the vector process $v = [v_1 \cdots v_L]^\top$ is modeled as a stationary stochastic process, according to $v(t) = H^0(q)e^0(t)$, $H^0 \in \mathbb{R}^{L\times p}(z)$, a proper and stable transfer function matrix with $p \leq L$ and $e := [e_1 \cdots e_p]^\top$ a vector white noise process with $var(e_i) = \sigma_i^2$. In this paper, we make the particular assumption that each node is directly influenced by only one innovation source, i.e. each column of $H^0(q)$ has only one nonzero entry. This encompasses the situation of full-rank uncorrelated disturbance ($p = L$, H^0 diagonal), as well as the situation of reduced-rank disturbance ($p < L$), see e.g. Weerts et al. (2018);
- r_j is an external excitation signal that is quasi-stationary and can be manipulated by users and is assumed to be independent of the disturbance v_j.

Further we assume that the network is well-posed and stable (Van den Hof et al., 2013) and that the standard regularity conditions on the data are satisfied, that are required for convergence of the prediction error identification method [1].

2.2 Topology information

To characterize the network interconnection structure (topology), we utilize a directed graph that captures both the locations and causal directions of interconnections in the network. This graph can be mathematically represented by a binary (adjacency) matrices $\mathcal{T}_G \in \mathbb{R}^{L\times L}$ and $\mathcal{T}_H \in \mathbb{R}^{L\times p}$, according to:

$$\mathcal{T}_G(j,i) = 0, \text{ if } G^0_{ji} \equiv 0; \quad \mathcal{T}_G(j,i) = 1, \text{ elsewhere},$$

$$\mathcal{T}_H(j,i) = 0, \text{ if } H^0_{ji} \equiv 0; \quad \mathcal{T}_H(j,i) = 1, \text{ elsewhere}.$$

In this study, $\mathcal{T}_G$ is assumed to be available, while $\mathcal{T}_H$ may be known or unknown. To streamline the utilization of information of the $\mathcal{T}_G$ and $\mathcal{T}_H$, we further introduce the subsequent set definitions.

Definition 1. In the graph of the network, we define the following sets of node indices for each node w_i:

- *$\mathcal{C}_i$: the set of node indices k, including i, for which either a directed path exists from w_k to w_i, or from w_i to w_k with $k \neq i$, or both w_k and w_i have a path from the same innovation source. The set $\mathcal{C}_i$ encompasses all node signals that are correlated with w_i;*
- *$\mathcal{J}_i$: the set of node indices k, including i, for which a directed path exists from w_i to w_k. Notably, set $\mathcal{J}_i$ is a subset of $\mathcal{C}_i$ by definition, i.e. $\mathcal{J}_i \subset \mathcal{C}_i$;*
- *$\mathcal{V}_j$: the set of node indices k, for which a path exists from the innovation source e of w_j.*

Considering the MISO subnetwork with the output node w_5 in Fig. 1, we have $\mathcal{N}_5 = \{1,2,3,4\}$ and $\mathcal{V}_5 = \{3,4,5\}$. For the in-neighbor nodes, we have $\mathcal{C}_1 = \{1,2,3,5\}$, $\mathcal{J}_1 = \{1,3,5\}$, $\mathcal{C}_2 = \mathcal{J}_2 = \{1,2,3,5\}$, $\mathcal{C}_3 = \{1,2,3,4,5\}$, $\mathcal{J}_3 = \{3,5\}$, $\mathcal{C}_4 = \mathcal{J}_4 = \{3,4,5\}$.

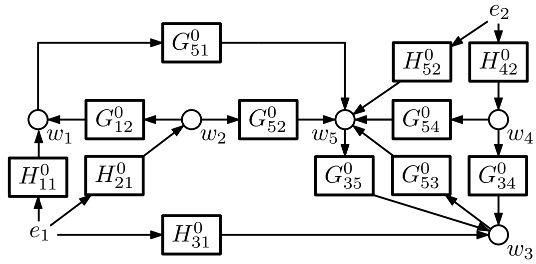

Fig. 1. An example dynamic network with five nodes.

2.3 One-step-ahead prediction error

The model validation technique in this study is based on the analysis of the prediction error on the output node $w_j(t)$ of the target MISO subnetwork, which is generated from the one-step-ahead predictor $\hat{w}_j(t \mid t-1)$ defined in Van den Hof et al. (2013). We consider a data-generating system denoted by $(G^0(q), H^0(q))$ and a specific model $(\hat{G}(q), \hat{H}(q))$, with the output prediction error:

$$\hat{\varepsilon}_j(t) = \hat{H}_j(q)^{-1}[w_j(t) - \sum_{k\in\mathcal{N}_j} \hat{G}_{jk}(q)w_k(t) - r_j(t)]. \qquad (2)$$

When substituting data from the data-generating system, this leads to the residual signal:

$$\hat{\varepsilon}_j(t) = \hat{H}_j(q)^{-1}\underbrace{(\sum_{k\in\mathcal{N}_j} \Delta\hat{G}_{jk}(q)w_k(t) + H^0_j(q)e_j(t))}_{\hat{v}_j(t)}, \qquad (3)$$

with $\Delta G_{jk}(q,\theta) = G^0_{jk}(q) - \hat{G}_{jk}(q)$. The signal $\hat{v}_j(t)$ is a reconstruction of the disturbance $v_j(t)$, provided that the model is correct.

3. LOCAL SUBNETWORK MODEL VALIDATION

The objective of the MISO subnetwork model validation is to determine whether the MISO model $\{\hat{G}_{j\mathcal{N}_j}(q), \hat{H}_j(q)\}$ can explain the measured data set $\{w_{\mathcal{N}_j}(t), w_j(t)\}$, potentially together with external input data $r_k(t)$, of the current system $\{G^0(q), H^0(q)\}$.

Typical residual tests in open-loop and closed-loop identification involve the autocorrelation function of the residual $\varepsilon(t)$ and its crosscorrelation with some given measured signals. In the network situation, we consider three types of correlation functions

- $R_{\hat{\varepsilon}_j}(\tau) := \bar{\mathbb{E}}\ \hat{\varepsilon}_j(t)\hat{\varepsilon}_j(t-\tau)$,
- $R_{\hat{\varepsilon}_j u_k}(\tau) := \bar{\mathbb{E}}\ \hat{\varepsilon}_j(t)u_k(t-\tau)$,
- $R_{\hat{v}_j u_k}(\tau) := \bar{\mathbb{E}}\ \hat{v}_j(t)u_k(t-\tau)$,

with $\bar{\mathbb{E}} := \lim_{N\to\infty} \frac{1}{N}\sum_{t=0}^{N-1} \mathbb{E}$ and $\mathbb{E}$ the expectation operator. $u_k(t)$ can be either a node signal $w_k(t)$, or an

[1] See Ljung (1999) page 249 and Söderström and Stoica (1989) Lemma B.3.

external signal $r_k(t)$. The following two null hypotheses form the test objectives of the correlation tests in the context of the dynamic network framework:

- ***Hypothesis*** $\mathcal{H}_a$***:*** *The residual* $\hat{\varepsilon}_j(t)$ *is a realization of a zero mean white noise process with variance* $\sigma^2_{e_j}$.
- ***Hypothesis*** $\mathcal{H}_b$***:*** *The residual* $\hat{\varepsilon}_j(t)$ *(or the predicted perturbation* $\hat{v}_j(t)$*) is independent of past input* u_k.

The validity of $\mathcal{H}_a$ and $\mathcal{H}_b$ can be checked using tests on the sample estimates of the three correlation functions.

Autocorrelation test Considering the residual $\{\hat{\varepsilon}_j(t)\}^N$, if $\mathcal{H}_a$ is true, it follows from a variant of the central limit theorem (Ljung, 1999) that the following distribution holds:

$$\frac{1}{\sqrt{N}}\sum_{t=1}^{N}\hat{\varepsilon}_j(t)\begin{bmatrix}\hat{\varepsilon}_j(t-1)\\ \vdots \\ \hat{\varepsilon}_j(t-M)\end{bmatrix} \sim As\ \mathcal{N}\left(0, \sigma^2_{e_j}\cdot I\right), \quad (4)$$

where M represents the considered number of time lags. For the time lag $\tau \in [1, M]$, the τ:th row of this vector is $\sqrt{N}\hat{R}^N_{\hat{\varepsilon}_j}(\tau)$, where $\hat{R}^N_{\hat{\varepsilon}_j}(\tau)$ is the estimated autocorrelation function defined as $\hat{R}^N_{\hat{\varepsilon}_j}(\tau) := \frac{1}{N}\sum_{t=1}^{N}\hat{\varepsilon}_j(t)\hat{\varepsilon}_j(t-\tau)$. The asymptotic normal distribution in (4) implies that the statistic $Q_a(N, M)$ defined as (Ljung (1999)):

$$Q_a(N,M) = \frac{N}{\left(\hat{R}^N_{\hat{\varepsilon}_j}(0)\right)^2}\sum_{\tau=1}^{M}\left(\hat{R}^N_{\hat{\varepsilon}_j}(\tau)\right)^2, \quad (5)$$

will have an asymptotic χ^2-distribution:

$$Q_a(N,M) \sim As\ \chi^2(M). \quad (6)$$

Consequently, hypothesis $\mathcal{H}_a$ can be tested using the asymptotic distribution in (6). Given a residual signal $\hat{\varepsilon}_j(t)$, the autocorrelation test for $\mathcal{H}_a$ is:

$$\begin{cases}\text{if } Q_a(N,M) \le c_\chi(\alpha, M), \text{ then accept } \mathcal{H}_a;\\ \text{otherwise, reject } \mathcal{H}_a \text{ with a risk equal to } \alpha,\end{cases} \quad (7)$$

where $c_\chi(\alpha, M)$ corresponds to the $1-\alpha$ quantile of the χ^2-distribution with M degrees of freedom, i.e. for $x \sim \chi^2(M)$ it follows that $\Pr(x \le c_\chi(\alpha, M)) = \alpha$.

Cross-correlation test Considering the residual $\{\hat{\varepsilon}_j(t)\}^N$ (or predicted disturbance $\{\hat{v}_j(t)\}^N$), if $\mathcal{H}_b$ is true, the statistic $\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k}$ defined as $\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k} =$

$$\frac{1}{N}\underbrace{\begin{bmatrix}u_k(1) & u_k(2) & \cdots & \cdots & u_k(N)\\ & \ddots & \cdots & \cdots & \vdots\\ & & u_k(1) & \cdots & u_k(N-M+1)\end{bmatrix}}_{P_{u_k}}\cdot\underbrace{\begin{bmatrix}\hat{\varepsilon}_j(1)\\ \vdots\\ \hat{\varepsilon}_j(N)\end{bmatrix}}_{\hat{\boldsymbol{\varepsilon}}_j}, \quad (8)$$

asymptotically converges to a zero-mean Gaussian distribution (Ljung (1999)):

$$\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k} \sim As\ \mathcal{N}(0, P), \text{ with } P = \frac{1}{N^2}P_{u_k}\Lambda_{\hat{\varepsilon}_j}P^\top_{u_k}, \quad (9)$$

with the residual auto-covariance matrix $\Lambda_{\hat{\varepsilon}_j} = \mathbb{E}\left[\hat{\boldsymbol{\varepsilon}}_j\hat{\boldsymbol{\varepsilon}}_j^\top\right]$ and the upper triangular matrix P_{u_k} as defined in (8). Douma et al. (2008) have shown that the statistic $\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k}$ will asymptotically converge to a χ^2 distribution with M degrees of freedom when scaled and squared:

$$Q_b(N,M) := \left[\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k}\right]^\top P^{-1}\left[\hat{\mathbf{R}}^N_{\hat{\varepsilon}_j u_k}\right] \sim As\ \chi^2(M). \quad (10)$$

Then given a residual $\hat{\varepsilon}_j(t)$ (or $\hat{v}_j(t)$) and an input signal $u_k(t)$, the cross-correlation test procedure for $\mathcal{H}_b$ is the same as (7) by replacing $Q_a(N, M)$ with $Q_b(N, M)$. If the noise model $\hat{H}_j(q)$ is not available, $\hat{v}_j(t)$ can be used instead of $\hat{\varepsilon}_j(t)$ for the cross-correlation test.

In the dynamic network framework, the result of the correlation test is usually influenced by multiple estimated transfer functions, mainly because the residual $\hat{\varepsilon}_j(t)$ and the predicted noise $\hat{v}_j(t)$ are generated from all the $\hat{G}_{j\mathcal{N}_j}(q)$ functions within a chosen MISO subnetwork. To specify the possible validation outcomes of a particular test, we introduce the concept of a target module set.

Definition 2. The target module set of a correlation test is defined as the set of all module transfer functions $G_{jk}(q)$ *that can be validated by this test. Dependent on whether one signal* $m(t)$ *or two signals* $m(t)$ *and* $n(t)$ *are used in the test, the corresponding target module set is denoted as* $\mathcal{S}_m$ *or* $\mathcal{S}_{mn}$*. If the correlation test is passed, all modules within* $\mathcal{S}_m$ *(or* $\mathcal{S}_{mn}$*) are assumed to be validated; consequently, the failed test implies the presence of at least one invalidated module within the target module set* $\mathcal{S}_m$ *(or* $\mathcal{S}_{mn}$*).*

For specifying the target module sets for each of the correlation tests, we will need to distinguish between three categories of prior noise information:

(a) $H^0_j(q)$ is available a priori;
(b) Only the topology $\mathcal{T}_H$ is available;
(c) No prior knowledge on $H^0_j(q)$ or $\mathcal{T}_H$ is available.

We will now specify the target module sets for the different correlations tests and for the different categories of prior noise information.

Proposition 1. The autocorrelation test using $\hat{\varepsilon}_j(t)$ *is applicable only with the noise information under situation (a). The target module set for the autocorrelation test is given by* $\mathcal{S}_{\hat{\varepsilon}_j} = \{G_{jk} \mid k \in \mathcal{N}_j\}$.

The autocorrelation test verifies whether $\hat{\varepsilon}_j(t)$ can be regarded as a white noise, which holds if and only if an accurate noise model $\hat{H}_j$ is available and $\Delta G_{jk}(q, \theta) = 0, \forall k \in \mathcal{N}_j$ in (3).

Proposition 2. The cross-correlation test using $\hat{\varepsilon}_j(t)$ *(or* $\hat{v}_j(t)$*) and* w_i *can be applied with the noise information in situations (a) and (b). In situation (a), all node signals* w_i *for* $i \in \mathcal{N}_j$ *can be used to conduct the test; In situation (b), node signals* w_i *for* $i \in \mathcal{N}_j \setminus \mathcal{V}_j$ *can be used to conduct the test. In both situations, the target module set of this test is given by* $\mathcal{S}_{\hat{\varepsilon}_j w_i} = \mathcal{S}_{\hat{v}_j w_i} = \{G_{jk} \mid k \in \mathcal{N}_j \cap \mathcal{C}_i\}$.

Given an accurate noise model, the correlation between $\hat{\varepsilon}_j$ and w_i is induced only by $\Delta G_{jk}(q, \theta) \neq 0$ in (3), where w_k has correlation to w_i. Without the noise model, $\hat{v}_j$ is correlated $w_i, i \in \mathcal{V}_j$ even with all $\Delta G_{jk}(q, \theta) = 0$, making such node signals unavailable to conduct the test.

Proposition 3. The cross-correlation test using $\hat{\varepsilon}_j(t)$ *(or* $\hat{v}_j(t)$ *and* r_i *can be used with the noise information under*

Table 1. Target module sets for different correlation tests and prior noise information categories.

Tests \ Target module sets / Noise information	(a) $H_j^0(q)$ & $\mathcal{T}_H$	(b) $\mathcal{T}_H$	(c) None
Autocorrelation test with $\hat{\varepsilon}_j$	$\mathcal{S}_{\hat{\varepsilon}_j}, \forall j$	$\emptyset$	$\emptyset$
Cross-correlation test with $\hat{\varepsilon}_j$ (or $\hat{v}_j$) and w_i	$\mathcal{S}_{\hat{\varepsilon}_j w_i}, i \in \mathcal{N}_j$	$\mathcal{S}_{\hat{v}_j w_i}, i \in \mathcal{N}_j \setminus \mathcal{V}_j$	$\emptyset$
Cross-correlation test with $\hat{\varepsilon}_j$ (or $\hat{v}_j$) and r_i	$\mathcal{S}_{\hat{\varepsilon}_j r_i}, i \in \mathcal{N}_j$	$\mathcal{S}_{\hat{v}_j r_i}, i \in \mathcal{N}_j$	$\mathcal{S}_{\hat{v}_j r_i}, i \in \mathcal{N}_j$

all situations (a), (b) and (c). The target module set of this test is given by $\mathcal{S}_{\hat{\varepsilon}_j r_i} = \mathcal{S}_{\hat{v}_j r_i} = \{G_{jk} \mid k \in \mathcal{N}_j \cap \mathcal{J}_i\}$.

For an excitation signal r_i independent of the disturbance, any correlation between $\hat{\varepsilon}_j$ and r_i will be induced by $\Delta G_{jk}(q,\theta) \neq 0$ in (3), where r_i has a path to G_{jk}.
The full proof of the above propositions is available in Shi (2023). The applicable scenarios and target module sets of three types of tests are summarized in Table 1. We further indicate tests that use signals already present in the data-generating system (w_k or any existing r_k) as passive tests. Active tests involve r signals that need to be added to the data generating system. All passive and active tests listed in Table 1 are regarded as applicable tests.

4. LOCAL FAULT DETECTION

In this section, we focus on detecting a fault in a MISO subnetwork with the output node w_j using model (in)validation techniques. A fault is characterized by a change of dynamics in any of the transfer functions $\{G_{jk}^0(q)\}, k \in \mathcal{N}_j$, and can include a transfer function becoming disconnected ($= 0$). Since the faulty module $G_{jk}^0(q)$ diverges from the healthy model $\hat{G}_{jk}(q)$, it can be detected using the correlation tests presented in Section 3.

During the fault detection (FD), it is presumed that the aim is to detect a potential fault within the target MISO subnetwork. For simplicity, we define the set of all modules in this subnetwork as U, where $U = \{G_{jk} | k \in \mathcal{N}_j\}$. It is not necessary to employ all tests listed in Table 1 for FD, instead, the objective is to choose the fewest tests from Table 1 such that their combined target module sets encompass set U. If any of the selected tests fails, we can conclude that a fault has been detected within the target MISO subnetwork. The selection of correlation tests for FD is based on the following proposition:

Proposition 4. For correlation tests in the dynamic network framework, it always holds that $\mathcal{S}_{\hat{\varepsilon}_j r_i} \subseteq \mathcal{S}_{\hat{\varepsilon}_j w_i} \subseteq \mathcal{S}_{\hat{\varepsilon}_j}$.

Proof: According to Definition 1, $\mathcal{J}_i \subseteq \mathcal{C}_i$. Therefore, $(\mathcal{N}_j \cap \mathcal{J}_i) \subseteq (\mathcal{N}_j \cap \mathcal{C}_i) \subseteq \mathcal{N}_j$. Consequently, $\{G_{jk} | k \in \mathcal{N}_j \cap \mathcal{J}_i\} \subseteq \{G_{jk} | k \in \mathcal{N}_j \cap \mathcal{C}_i\} \subseteq \{G_{jk} | k \in \mathcal{N}_j\}$. ■

Proposition 4 indicates that the target modulus sets for the three types of correlation tests listed in Table 1 gradually decrease in size. Consequently, for each category of prior noise information, it is important to prioritize the tests in the diagonal blocks of Table 1 when selecting test(s) for FD, since their target module sets can include the most modules within the set U.

5. LOCAL FAULT DIAGNOSIS

Upon detecting a fault in the target subnetwork, the subsequent step for fault diagnosis is to identify its root cause, which is to locate the specific faulty module. The process to determine the location of a detected fault is defined as fault isolation (FI) (Isermann (2006)). The FI task within the target subnetwork can also be addressed by the correlation test. Each correlation test has its own target module set, and the modules in each set can vary. To facilitate the description of all these target module sets, we introduce the following definition:

Definition 3. Given a target subnetwork with output node w_j, and a selected category of prior noise information, then the set $\mathfrak{S}_j$ is defined as the set of target module sets of all applicable tests.

To illustrate, the set $\mathfrak{S}_j$ collects all target module sets in a pre-chosen column of Table 1. According to Definition 2, the tests corresponding to the target module sets containing the faulty module will fail, while the tests corresponding to the target module sets not containing the faulty module will pass. This allows us to locate the detected faults in the target module set of a failed test or even on a specific module by combining the results from different tests.

Example 1. For the network shown in Fig. 1, there is a detected fault in the MISO subnetwork with the output node w_5. Given the prior noise information under situation (b), set $\mathfrak{S}_5$ contains the following target module sets:

1. $\mathcal{S}_{\hat{v}_5 w_1} = \{G_{5k} | k \in \mathcal{N}_5 \cap \mathcal{C}_1\} = \{G_{51}, G_{52}, G_{53}\}$,
2. $\mathcal{S}_{\hat{v}_5 r_1} = \{G_{5k} | k \in \mathcal{N}_5 \cap \mathcal{J}_1\} = \{G_{51}, G_{53}\}$,
3. $\mathcal{S}_{\hat{v}_5 w_2} = \{G_{5k} | k \in \mathcal{N}_5 \cap \mathcal{C}_2\} = \{G_{51}, G_{52}, G_{53}\}$,
4. $\mathcal{S}_{\hat{v}_5 r_2} = \{G_{5k} | k \in \mathcal{N}_5 \cap \mathcal{J}_2\} = \{G_{51}, G_{52}, G_{53}\}$,

Assuming that there is only one fault in this target subnetwork, we can try to isolate the fault on a certain module by combining the applicable tests. If the fault is on G_{52}^0, the first test will fail and the second test will pass. Given that their target module sets follow $\mathcal{S}_{\hat{v}_5 w_1} \setminus \mathcal{S}_{\hat{v}_5 r_1}$, the test results will indicate that the fault location is on G_{52}. If the fault is on G_{53}^0, the first and second tests will fail at the same time. Given that $\mathcal{S}_{\hat{\varepsilon}_5 w_1} \cap \mathcal{S}_{\hat{\varepsilon}_5 r_3} = \{G_{53}\}$, the test results indicate that the fault location is on G_{53}.

As can be seen from Example 1, smaller module sets can be obtained by performing set operations such as taking the intersection or difference of different target module sets. These smaller sets contain fewer modules or even one module after sufficient set operations. In the process of obtaining these smaller sets, the tests corresponding to the used target module sets will contribute to FI. The resulting smaller set is further specified by the following definition:

Definition 4. Define a set of modules $\mathbb{S}(G_{jk})$ as a set that (a) is composed of any element in set $\mathfrak{S}_j$, or the result of any sequence of set operations of intersection and difference applied to elements of $\mathfrak{S}_j$, (b) contains module G_{jk}, and (c) is of minimal cardinality.

Note that $\mathbb{S}(G_{jk})$ may be non-unique. If $\mathbb{S}(G_{jk})$ only contains module G_{jk}, a fault occurring on module G_{jk} can be isolated; if set $\mathbb{S}(G_{jk})$ contains other modules apart from G_{jk}, a fault occurring on module G_{jk} can only be isolated in the range of set $\mathbb{S}(G_{jk})$. This analysis procedure can also indicate the contributing correlation tests for FI, i.e. the tests that are used to obtain $\mathbb{S}(G_{jk})$ can guarantee to isolate a fault within the range of $\mathbb{S}(G_{jk})$. Since the analysis on $\mathbb{S}(G_{jk})$ fully relies on set $\mathfrak{S}_j$, it can be done before the data-collecting experiment. The following strategy will be followed to generate set $\mathbb{S}(G_{jk})$ for each G_{jk} in set U:

1. The algorithm starts by initializing $\mathbb{S}(G_{jk})$ for each G_{jk} in set U to be U, i.e. $\mathbb{S}(G_{jk}) = U$;
2. For each pair of sets $\mathcal{S}_{pq}$, $\mathcal{S}_{mn}$ in $\mathfrak{S}_j$, compute their intersection $\mathcal{S}_{pq} \cap \mathcal{S}_{mn}$ or difference $\mathcal{S}_{pq} \setminus \mathcal{S}_{mn}$ in one loop. Mark the result as a module set $\mathcal{S}_{new}$;
3. For the resulting set $\mathcal{S}_{new}$ that satisfies $\mathcal{S}_{new} \neq \emptyset$ or $\mathcal{S}_{new} \notin \mathfrak{S}_j$, add $\mathcal{S}_{new}$ to $\mathfrak{S}_j$;
4. For each module $G_{jk} \in \mathcal{S}_{new}$, if the size of $\mathcal{S}_{new}$ is smaller than $\mathbb{S}(G_{jk})$, update $\mathbb{S}(G_{jk})$ to be $\mathcal{S}_{new}$;
5. Record the correlation tests corresponding to the sets used to obtain the new $\mathbb{S}(G_{jk})$ for module G_{jk};
6. Loop from step 2 to step 5 till no smaller sets $\mathcal{S}_{new}$ can be obtained.

Step 5 of the above algorithm records all the contributing correlation tests for FI. If the contributing tests include active tests, the additionally required r signals need to be allocated to the system before collecting data.

6. NUMERICAL ILLUSTRATION

A numerical illustration of the proposed FDD procedure is provided. The data-generating network depicted in Fig. 2 is used. Disturbances v_1, v_2, v_3 on each node originate from independent innovation sources, detailed as:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} H_{11}^0 & 0 & 0 \\ 0 & H_{22}^0 & 0 \\ 0 & 0 & H_{33}^0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}, \quad (11)$$

where $Var(e_1) = 0.1, Var(e_2) = 0.2, Var(e_3) = 0.3$. A potential fault is considered to occur on module G_{12}^0 when it changes to a faulty module G_{12}^f. Specifications regarding the module transfer functions and noise models can be found in the appendix. FD and FI are conducted for three scenarios, each with different levels of noise information under situations (a), (b), and (c) as illustrated in Table 1. Both FD and FI consist of two phases: the correlation test selection and the data-collecting experiments, which will be illustrated sequentially. Based on Definition 1, we have the node index sets $\mathcal{N}_1 = \{2, 3\}$, $\mathcal{V}_1 = \{1\}$, $\mathcal{C}_2 = \mathcal{J}_2 = \{1, 2, 3\}$, $\mathcal{C}_3 = \{1, 2, 3\}$, $\mathcal{J}_3 = \{1, 3\}$, which are used to determine the target module sets of all tests.

Fault detection The task of FD is to detect if there is a fault in $U = \{G_{12}, G_{13}\}$. During the test selection phase, the autocorrelation test using residual $\hat{\varepsilon}_1$ is chosen for situation (a) since its target module set $\mathcal{S}_{\hat{\varepsilon}_1} = \{G_{1k} | k \in \mathcal{N}_1\} = \{G_{12}, G_{13}\}$ can already cover set U without relying on excitation signals. For situation (b), the cross-correlation test using $\hat{v}_1$ and w_2 with the target module set $\mathcal{S}_{\hat{v}_1 w_2} = \{G_{1k} | k \in \mathcal{N}_1 \cap \mathcal{C}_2\} = \{G_{12}, G_{13}\}$ is selected for the same reason as in situation (a). An alternative could be another cross-correlation test employing $\hat{v}_1$ and w_3, given that $\mathcal{S}_{\hat{v}_1 w_3} = \{G_{1k} | k \in \mathcal{N}_1 \cap \mathcal{C}_3\} = \{G_{12}, G_{13}\}$. For situation (c), the cross-correlation test that uses $\hat{v}_1$ and r_2 is chosen since $\mathcal{S}_{\hat{v}_1 r_2} = \{G_{1k} | k \in \mathcal{N}_1 \cap \mathcal{J}_2\} = \{G_{12}, G_{13}\} = U$. This is an active test given that r_2 is absent from the data-generating network (see Fig. 2). Therefore before the subsequent data-collecting phase, r_2 is added to the system for situation (c) as a white noise signal, characterized by $Var(r_2) = 5$.

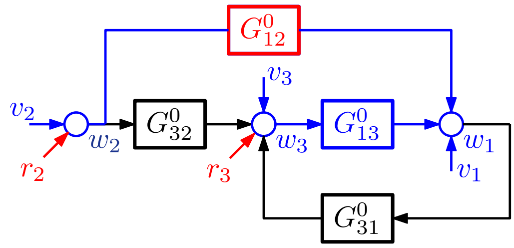

Fig. 2. Example of a 3-node network, the target MISO subnetwork is marked in blue. The potential faulty module G_{12} is marked in red. The excitation r_2 and r_3 (red) are not originally presented in the data-generating system but can be added for FD and FI.

During the data-collecting phase, the selected tests for FD are employed for each situation. The simulation runs with a total data length of $N = 10000$. Initially, the module G_{12}^0 is configured to a healthy state, and it transitions to the faulty state G_{12}^f after $N = 5001$. Test levels are computed at intervals of 500 steps with the collected data. Each time before $N \leq 5000$ when computing the test level, all measured data is used; after $N > 5000$, the test level is computed only based on the newest 5000 data point. The test result of situation (a) is given in Fig. 3, while situations (b) and (c) show similar results. The data-collecting experiment incorporated 100 Monte Carlo simulations, the center of each sample point in Fig. 3 denotes the mean value of 100 simulations, while the vertical line indicates its standard deviation.

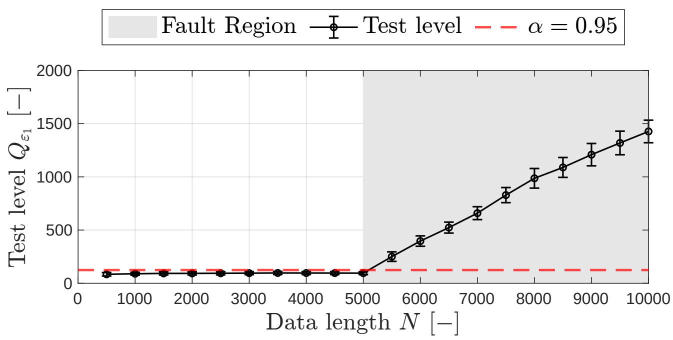

Fig. 3. Detect a fault with the validation tests in the target MISO subnetwork as in Fig. 2 for noise information (a) with test level $Q_{\hat{\varepsilon}_1}$ (black) against the confidence threshold with $\alpha = 0.95$ (red).

As shown in Fig. 3, a fault within the local subnetwork is accurately detected.

Fault isolation The task of FI is to determine where the fault originated. During the test selection phase, the sets $\mathfrak{S}_1$ are derived, collecting all target modules listed in Table 1 for each situation, which is the input of the algorithm in Section 5. For FI under each situation, the algorithm selects two tests. Test 1 is the test used for FD with its target module set equal to U, and Test 2 is a cross-correlation test using r_3 with its target module set $\mathcal{S}_{\hat{v}_1 r_3} = \{G_{1k} | k \in \mathcal{N}_1 \cap \mathcal{J}_3\} = \{G_{13}\}$. Because the combination of these two tests leads to $\mathbb{S}(G_{12}) = U \setminus \mathcal{S}_{\hat{v}_1 r_3} = \{G_{12}\}$ and $\mathbb{S}(G_{13}) = \mathcal{S}_{\hat{v}_1 r_3} = \{G_{13}\}$. Test 2 is an active test so r_3 is added to the system before the subsequent data-collecting phase for each situation.

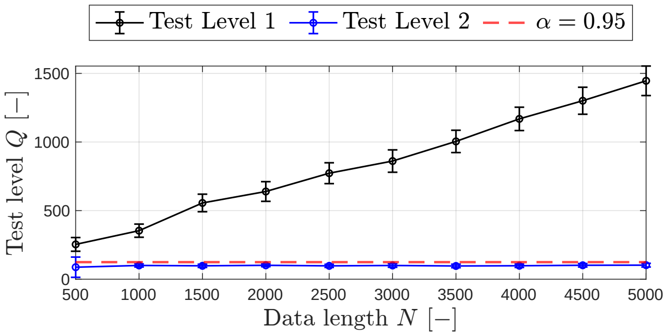

Fig. 4. Isolate a fault with the validation tests in the target MISO subnetwork as in Fig. 2 for noise information (a) with the test levels $Q_{\hat{\varepsilon}_1}$ (black) and $Q_{\hat{\varepsilon}_1 r_3}$ (blue).

For the selected tests, the data-collecting phase for FI is conducted with length $N = 5000$, with test levels being calculated every 500 steps and each time using all measured data. The test results of situation (a) are illustrated in Fig. 3, while situations (b) and (c) show similar results. The result shows a rapid increase in the test level of Test 1 (black), surpassing the confidence bound, while the test level of Test 2 (blue) consistently stays below the bound. This outcome indicates that Test 1 fails and Test 2 is passed. Combining two test results, it is deduced that the detected fault is isolated in module G_{12}.
In the FDD procedure, the active test using r_3 is required for FI under all situations (a), (b), and (c), while the active test using r_2 is additionally required only for situation (c). Therefore, the numerical illustration for different levels of noise information underscores the inherent trade-off between noise information and the number of potentially required active tests for FDD.

7. CONCLUSION

A local model-based FDD method has been developed for interconnected systems, which employs model invalidation tests in a dynamic network framework. Existing auto- and cross-correlation tests are generalized to the dynamic network situation for local subnetwork model validation. Based on different levels of prior noise information, available correlation test(s) can be selected for FD within a local subnetwork. Subsequently, for the fault diagnosis, it is demonstrated that combining results from multiple correlation tests enables the isolation of the detected fault more precisely, even narrowing it down to a specific module.

Appendix A. SIMULATED NETWORK

For the numerical illustration, the following transfer functions for the data-generating network as in Fig. 2 are used.

$$G_{13}^0(q) = \frac{-0.2q^{-1}}{1-1.3q^{-1}+0.6q^{-2}}, \quad H_{11}^0(q) = \frac{1+0.52q^{-1}}{1+0.41q^{-1}},$$
$$G_{31}^0(q) = \frac{0.39q^{-1}}{1-0.8q^{-1}+0.2q^{-2}}, \quad H_{22}^0(q) = \frac{1+0.44q^{-1}}{1+0.35q^{-1}},$$
$$G_{32}^0(q) = \frac{-0.3q^{-1}}{1+0.6q^{-1}+0.2q^{-2}}, \quad H_{33}^0(q) = \frac{1+0.52q^{-1}}{1+0.45q^{-1}}.$$
$$G_{12}^0(q) = \frac{0.39q^{-1}}{1-0.8q^{-1}+0.2q^{-2}}, \quad G_{12}^f(q) = \frac{0.312q^{-1}}{1-0.8q^{-1}+0.2q^{-2}}$$

REFERENCES

Dankers, A.G., Van den Hof, P.M.J., Bombois, X., and Heuberger, P.S.C. (2016). Identification of Dynamic Models in Complex Networks With Prediction Error Methods: Predictor Input Selection. *IEEE Trans. on Automatic Control*, 61(4), 937–952.

Douma, S.G., Bombois, X., and Van den Hof, P.M.J. (2008). Validity of the standard cross-correlation test for model structure validation. *Automatica*, 44, 1285–1294.

Dowdeswell, B., Sinha, R., and MacDonell, S.G. (2020). Finding faults: A scoping study of fault diagnostics for Industrial Cyber–Physical Systems. *Journal of Systems and Software*, 168, 110638.

Fonken, S.J.M., Ramaswamy, K.R., and Van den Hof, P.M.J. (2022). A scalable multi-step least squares method for network identification with unknown disturbance topology. *Automatica*, 141, 110295.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. Marcel Dekker, Boca Raton, FL.

Gevers, M., Bazanella, A.S., and Pimentel, G.A. (2019). Identifiability of dynamical networks with singular noise spectra. *IEEE Trans. Aut. Control*, 64(6), 2473–2479.

Gonçalves, J. and Warnick, S. (2008). Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *IEEE Tr. Aut. Con.*, 53(7), 1670–1674.

Isermann, R. (2006). *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer, Berlin, Heidelberg.

Khorasgani, H., Jung, D., and Biswas, G. (2015). Structural Approach for Distributed Fault Detection and Isolation. *IFAC-PapersOnLine*, 48(21), 72–77.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.

Materassi, D. and Salapaka, M. (2012). On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Trans. Automatic Control*, 57(7), 1765–1777.

Ramaswamy, K.R., Csurcsia, P.Z., Schoukens, J., and Van den Hof, P.M. (2022). A frequency domain approach for local module identification in dynamic networks. *Automatica*, 142, 110370.

Shi, Y. (2023). *Fault detection and diagnosis using the dynamic network framework*. MSc thesis, Eindhoven University of Technology.

Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice-Hall, New York, NY.

Van den Hof, P.M.J., Dankers, A.G., Heuberger, P.S.C., and Bombois, X. (2013). Identification of dynamic models in complex networks with prediction error methods - basic methods for consistent module estimates. *Automatica*, 49(10), 2994–3006.

Weerts, H.H.M., Van den Hof, P.M.J., and Dankers, A.G. (2018). Prediction error identification of linear dynamic networks with rank-reduced noise. *Automatica*, 98, 256–268.